



UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Volveau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 1447

Programme 3

*Intelligence artificielle, Systèmes cognitifs et
Interaction homme-machine*

SYSTEMES MARKOVIENS DISCRETS STATIONNAIRES ET APPLICATIONS

Jean PELLAUMAIL

Juin 1991



★ R R . 1 4 4 7 ★

Systèmes markoviens discrets stationnaires et applications

Discrete stationary Markov processes for applications

Jean PELLAUMAIL

IRISA - Publication Interne n° 579 - mars 1991

Programme 3

Campus Universitaire de Beaulieu
35042 - RENNES CEDEX
FRANCE
Téléphone : 99.36.20.00
Télex : UNIRISA 950 473F
Télécopie : 99.38.38.32

Systèmes markoviens discrets stationnaires et applications

Discrete stationary Markov processes for applications

Jean PELLAUMAIL

I.N.S.A.
20, avenue des Buttes de Coësmes
35043 RENNES CEDEX
FRANCE

Résumé

Ce travail (270 pages) se propose essentiellement de faire le point sur les résultats mathématiques actuellement effectivement numériquement opérationnels pour le calcul exact des probabilités stationnaires d'un système markovien.

Abstract

This work (270 pages) is a survey on the mathematical results which are needed for computing the steady-state probabilities of discrete Markov processes.

INTRODUCTION

Ce cours s'adresse d'abord aux mathématiciens, aux ingénieurs et aux futurs ingénieurs, spécialisés dans l'étude et la modélisation de phénomènes dont l'équilibre global est assuré par un grand nombre de transitions élémentaires. Cette situation apparaît dans des domaines très variés : informatique, téléphonie, fiabilité et sûreté de fonctionnement, économie, biologie, etc...

En général, placé dans un tel contexte, un ingénieur commence par étudier son problème à l'aide de la simulation. De ce point de vue, le **chapitre 3 peut être considéré comme une introduction à l'ensemble de ce cours**. Il situe la simulation en général et la simulation informatique en particulier par rapport au reste du cours. Il contient "en germe" l'essentiel des notions théoriques générales dont la formalisation mathématique est précisée dans les autres chapitres : notamment, l'accent y est mis sur la notion indispensable d'"évolution markovienne". **Un simulateur est toujours markovien**.

Les hypothèses sont précisées au chapitre 1. On y montre l'existence et l'unicité d'une solution qui donne l'évolution théorique exacte au cours du temps. Sauf dans quelques cas très particuliers qui sont étudiés au chapitre 8, cette solution est peu exploitable numériquement. Par contre, les équations en régime stationnaire, c'est à dire en situation d'équilibre global, sont "relativement" simples : ces équations et l'unicité de leur solution sont établies au chapitre 2 sous des hypothèses très générales.

Ces chapitres 1 et 2 ne prétendent évidemment pas résumer les travaux innombrables et prestigieux portant sur les processus markoviens. Par contre, comme l'ensemble de ce cours, ils sont "self-contained", c'est à dire qu'ils ne font pas appel à d'autres connaissances spécialisées, ils donnent la base théorique minimale suffisante pour supporter les autres chapitres et toutes les démonstrations y sont "complètes".

Les chapitres 4 à 9 se proposent de faire le point sur les résultats mathématiques actuellement effectivement numériquement opérationnels pour le calcul **exact** en "situation stationnaire". Dans ce domaine, les résultats les plus spectaculaires sont associés à la "forme produit" (théorème de Jackson, BCMP, etc...) : ils sont établis au chapitre 4. Par ailleurs, les méthodes classiques de résolution des systèmes linéaires, notamment quand interviennent des matrices creuses, doivent servir de référence : les chapitres 5 et 6 situent ces méthodes par rapport au contexte spécifique des systèmes en équilibre. Les équations aux dérivées partielles

linéaires d'une part et des matrices de Léontief - bien connues en économie - d'autre part sont proposées en exemple au chapitre 6.

La théorie des graphes a toujours été un intermédiaire technique irremplaçable dans l'étude des systèmes markoviens discrets ; cet aspect est étudié au chapitre 7 dans une perspective tout à fait nouvelle et cruciale relativement à l'ensemble de ce cours.

Le chapitre 9 donne l'essentiel des résultats liés à l'utilisation de la fonction génératrice : cet outil est particulièrement efficace en biologie.

Dans les chapitres 1 à 9, l'ensemble des états est supposé fini, ou "presque" fini : c'est le cas discret. De plus, les hypothèses y sont suffisamment fortes pour qu'il ne soit pas nécessaire d'utiliser explicitement la notion de "temps d'arrêt". Pour le lecteur qui souhaite dépasser ce stade, une ouverture sur la théorie mathématique générale est proposée au chapitre 10. Quelques commentaires sur le cas dénombrable sont proposés en 3.A.6 et 7.D.6.

Au niveau de la présentation, on impose toujours $a_{i,j} = a(i,j)$, $b_{u,v} = b(u,v)$, etc... Le choix entre ces deux notations est purement typographique.

Ce cours est un cours de mathématiques orienté vers les applications. Chaque chapitre est ordonné autour d'une (ou plusieurs) technique mathématique et non pas relativement à un type d'usager. Par contre, seuls les techniques actuellement utilisables sont étudiées. Pour la commodité du lecteur pressé - qui ne l'est pas ? - chaque chapitre peut être lu presque indépendamment des autres. Ceci conduit à quelques répétitions. La liaison entre les divers chapitres est assurée par de nombreux renvois. De plus, un index est donné à la fin du cours, après les références. Enfin, pour l'essentiel, ce cours est original, soit quant au fond, soit quant à la forme.

Je tiens à remercier Madame Martinez qui a assuré la frappe et la mise en page de ce travail et qui a effectué, avec beaucoup de patience, de multiples corrections. Je remercie aussi M. Lebah, P. Leguesdron et L.M. Le Ny pour leurs remarques pertinentes.

TABLE DES MATIERES

CHAPITRE 1 : Processus markoviens discrets

A Exemples de base : 1. Connaissances requises. 2. Loi binomiale et marche aléatoire. 3. Loi et processus de Poisson. 4. Loi exponentielle.

B Loi d'un processus : 1. Processus stochastique. 2. Loi d'un processus. 3. Processus markovien. 4. Semi-groupe de transition.

C Chaîne de Markov : 1. Matrice de transition. 2. Semi-groupe de transition. 3. Matrice stochastique. 4. Pile ou face. 5. Une stratégie élémentaire. 6. Stock à point de commande.

D Processus markovien à temps continu : 1. Matrice d'évolution. 2. Equations de Chapman-Kolmogorov. 3. Equations backward de Kolmogorov. 4. Système différentiel linéaire. 5. Semi-groupe de transition.

E Exemples : 1. Loi exponentielle. 2. Loi d'Erlang. 3. Processus de Poisson. 4. File M/M/1. 5. File E2/M/1. 6. Automate exponentiel.

F Etats fictifs : 1. Processus non markovien. 2. Loi PH. 3. Loi hyper-exponentielle. 4. Modélisation avec pivot. 5. Modélisation de Cox "positive". 6. Loi exponentielle déguisée.

G Approximation de type CP : 1. Introduction. 2. Hypothèses et notations. 3. Lemme préliminaire. 4. Théorème d'approximation. 5. Remarques. 6. Loi de type PH. 7. Lemme sur les lois exponentielles. 8. Théorème d'équivalence. 9. Remarques. 10. Contre-exemples.

CHAPITRE 2 : Régime stationnaire

A Coefficient de contraction : 1. Préliminaires. 2. Ecart angulaire. 3. Coefficient de contraction. 4. Définition de μ . 5. Matrices stochastiques.

B Produit de matrices positives : 1. Vecteur propre. 2. Matrices variables. 3. Matrice irréductible.

C Modélisation et régime stationnaire : 1. Introduction. 2. Exemples et remarques. 3. Cas d'une chaîne homogène. 4. Cas à temps continu.

D Les équations d'équilibre : 1. Hypothèses de base. 2. Cas des proces-

sus. 3. Cas fini irréductible. 4. Résolution de $q_d = q_a$. 5. Théorème ergodique. 6. Processus stationnaire.

E Théorème ergodique ponctuel : 1. Introduction. 2. Exemple de base. 3. Dédoublage de processus. 4. Variables de comptage.

F Etude trajectorielle : 1. Introduction. 2. Trajectoire. 3. Fonction presque markovienne. 4. Formule de Little.

CHAPITRE 3 : Simulation

A Modélisation : 1. Introduction. 2. Cadre markovien. 3. Arrivées poissonniennes. 4. Répétition d'appels et caractère markovien. 5. Régime stationnaire. 6. Cas E infini.

B Nombres au hasard : 1. Définition mathématique. 2. Construction. 3. Sondages et codages. 4. Tests. 5. Générateur et simulation.

C Simulation expérimentale : 1. Un simulateur est markovien. 2. E fini. 3. Modélisation d'un simulateur.

D Simulation informatique d'une chaîne de Markov : 1. Préliminaires. 2. Un jeu élémentaire. 3. Bus avec blocages. 4. Les transitions. 5. Le déroulement du programme. 6. Interprétation mathématique. 7. Viabilité de la méthode.

E Processus évoluant à temps continu : 1. Cas d'un processus markovien homogène. 2. Discrétisation du temps. 3. File GI/GI/1. 4. Temps d'arrêt.

F Le progiciel QNAP2 : 1. Langages de simulation. 2. Les déclarations dans QNAP2. 3. Le type queue. 4. Sémaphore, resource, flag. 5. Résolution. 6. Autres aspects.

G Conclusion.

CHAPITRE 4 : Solution à forme produit

A Serveur central : 1. Réseau de files d'attente. 2. Notations générales. 3. Hypothèses du serveur central. 4. Probabilité stationnaire. 5. Remarques. 6. Routages fixes. 7. Routages variables.

B Routages fixes, monoclasse : 1. Hypothèses : cas fermé. 2. Théorème de Jackson. 3. Cas ouvert. 4. La constante de normalisation. 5. Sensibilité.

C Routages fixes, multiclasse : 1. Hypothèses : cas fermé. 2. Théorème

BCMP (macro-états). 3. Remarques.

D Modification des taux de service : 1. Introduction. 2. Hypothèses de base. 3. Théorème de stabilité. 4. Conservation de la balance locale.

E Produit de deux réseaux : 1. Introduction. 2. Hypothèses. 3. Théorème de composition. 4. Conservation de la balance locale.

F Exemples : 1. Un seul client et corollaires. 2. Premier arrivé, premier servi.

CHAPITRE 5 : Matrices

A Analyse numérique : 1. Notations matricielles. 2. Hypothèses générales. 3. Remarques générales.

B Matrices tridiagonales : 1. Cadre général. 2. La factorisation $t = t't''$. 3. Cas limite. 4. Cas général. 5. Algorithme.

C Matrices tridiagonales par blocs : 1. Hypothèses. 2. Méthode générale de résolution. 3. Variables auxiliaires. 4. Retour sur la factorisation.

D La méthode des convexes : 1. Cadre général. 2. Les convexes $C(k)$. 3. Mise en oeuvre. 4. Convergence. 5. Un cadre plus précis.

E Un cas particulier : 1. Introduction. 2. Notations et hypothèses générales. 3. Théorème. 4. Convergence des convexes $C(k)$. 5. Exemples et remarques. 6. Généralisation. 7. Conclusion.

CHAPITRE 6 : Splitting, chaîne incluse et exemples

A Splitting : 1. Introduction. 2. Théorème. 3. Remarques. 4. Inversion de $(d-c)$. 5. Cas tridiagonal par blocs.

B Cas avec pivot : 1. Introduction. 2. Théorème. 3. Remarques. 4. Cas tridiagonal par blocs.

C Chaîne incluse : 1. Introduction. 2. Inversion de $(d-c)$. 3. Probabilité stationnaire pour une trajectoire. 4. Transition H . 5. Processus inclus. 6. Chaîne incluse.

D Points absorbants : 1. Introduction. 2. Probabilités stationnaires. 3. Fiabilité et sûreté de fonctionnement. 4. Pénalisation.

E Equations aux dérivées partielles : 1. Introduction. 2. Les différences

finies. 3. Equation de type parabolique. 4. Equation de type elliptique ou hyperbolique. 5. Les conditions aux limites. 6. Les analogies. 7. Les différences. 8. Cas avec second membre.

F Matrices de Léontief : 1. Introduction. 2. Equilibre des prix. 3. Equilibre des flux économiques. 4. Aspect descriptif. 5. Sensibilité.

G Conclusion.

CHAPITRE 7 : Graphes

A Réseaux fluviaux : 1. Graphes et matrices d'évolution. 2. Vocabulaire et définition. 3. Réseau fluvial. 4. Cas avec second membre. 5. Irrigabilité. 6. M-matrice inversible.

B Relations de base : 1. Notation β . 2. Notation $\gamma(t,u,v,w)$. 3. Notation $\gamma'(u,t,v,w)$. 4. Notation $\beta'(v,u,w)$. 5. Théorème.

C Cas avec second membre : 1. Théorème fondamental. 2. Cas $x'=0$: précision. 3. Lien avec les déterminants.

D Cas tridiagonal par blocs : 1. Introduction. 2. Notations W et $B(e,w)$. 3. Théorème. 4. Commentaires heuristiques. 5. Remarques techniques. 6. Cas E infini.

E Liens entre homogène et non homogène : 1. Introduction. 2. Drainage. 3. Théorème de base. 4. Autre présentation. 5. Remarques.

CHAPITRE 8 : Régime transitoire

A Exponentielle de matrice : 1. Introduction. 2. Propriétés élémentaires. 3. Cas commutatif. 4. Cas positif. 5. Calcul de e^{mt} . 6. Condition initiale. 7. Remarques.

B Splitting et chaîne incluse : 1. Exponentielle de $(m+n)$. 2. Remarques. 3. Cadre probabiliste. 4. Interprétation heuristique. 5. Remarque. 6. Cas triangulaire par blocs.

C Flots poissonniens : 1. Proposition. 2. Cadre probabiliste. 3. Balance locale. 4. Remarques.

D Processus réversibles : 1. Introduction. 2. Pseudo-commutativité. 3. Processus réversible. 4. Exemples. 5. Renversement du temps.

E Processus de renouvellement : 1. Définition. 2. Aspect théorique.

3. Simulation. **4.** Aspect technique.

F Transitoire stationnaire et fiabilité : **1.** Introduction. **2.** Cadre probabiliste. **3.** Durée de vie. **4.** Chaîne incluse. **5.** Exemple de base. **6.** Un cas tridiagonal par blocs. **7.** File unique. **8.** Remarques.

CHAPITRE 9 : Fonction génératrice

A Notions de base : **1.** Introduction. **2.** Cas à une dimension. **3.** Développement en $z = 1$. **4.** Cas multidimensionnel. **5.** Constante de normalisation.

B File M/PH/1 : **1.** Introduction. **2.** Hypothèses mathématiques. **3.** Chaîne incluse. **4.** Régime stationnaire. **5.** Calcul de H. **6.** Calcul des moments.

C Processus de naissances et de morts : **1.** Hypothèses. **2.** Equations de Chapman-Kolmogorov. **3.** Fonction génératrice. **4.** Propriétés générales de la solution. **5.** Solution associée à $z=1$.

D Cas presque linéaire en z : **1.** Hypothèses. **2.** Equations de Chapman-Kolmogorov. **3.** Résolution. **4.** Calcul des moments. **5.** Exemples.

E File M/M/ ∞ : **1.** Hypothèses. **2.** Résolution générale. **3.** Solution particulière. **4.** Calcul final.

F Etude abstraite : **1.** Préliminaire. **2.** Equations. **3.** Résolution.

CHAPITRE 10 : Introduction à la théorie

A Préliminaires : **1.** Objet du chapitre. **2.** Rigueur mathématique. **3.** L'axiome du choix. **4.** Loi de probabilité. **5.** Espace fondamental.

B Processus canonique : **1.** Loi d'un processus. **2.** Le théorème de Kolmogorov-Bochner. **3.** Exemple. **4.** Régime stationnaire. **5.** Automate exponentiel stationnaire. **6.** Processus de Poisson.

C Processus cadlag et temps d'arrêt : **1.** Introduction. **2.** Modification d'un processus. **3.** Processus indistingables. **4.** Processus cadlag. **5.** Existence de modification cadlag. **6.** Filtration. **7.** Temps d'arrêt.

D Constructions poissonniennes : **1.** Introduction. **2.** Cas E fini et $T = \mathbb{R}^+$. **3.** Construction de X. **4.** Remarques. **5.** Les limites de telles constructions.

E Processus markovien stationnaire : 1. Processus markovien. 2. Processus stationnaire. 3. Chaîne incluse. 4. Intégrale stochastique.

F Conclusion.

Chapitre 1

Processus markoviens discrets**A. Exemples de base****A.1 Connaissances requises**

Pour la plus grande partie de ce cours, les connaissances requises se limitent à quelques notions de base en **théorie des matrices** et en **probabilité**. Plus précisément, on supposera que le lecteur est familiarisé avec les mots ou expressions qui suivent : vecteur, matrice, norme, système d'équations linéaires, loi de probabilité, variable aléatoire, σ -additivité, probabilité conditionnelle $p(A|B)$ (qui se lit probabilité de A sachant B), formule de Bayes. Les lois les plus usuelles - loi binomiale, loi de Poisson, loi exponentielle, loi d'Erlang - seront rappelées, au cours de ce paragraphe A ou au paragraphe E, à l'occasion d'exemples. La transformée de Laplace sera utilisée au paragraphe 1.F.

Un bref rappel sur quelques notions matricielles est donné en 2.A.1. Si g est une matrice, $g_{i,j}$ désigne le terme de la i -ième ligne et de la j -ième colonne de cette matrice g . La notion de déterminant sera évoquée au chapitre 7 et celle de vecteur propre en 2.B.

A.2 Loi binomiale et marche aléatoire

a) Paul joue à pile ou face ; si la pièce retombe sur pile, il gagne un franc ; si elle retombe sur face, il perd un franc ; soit X_k ses gains ou pertes (globalement) après la k -ième partie. On dit que $(X_k)_{k \in \mathbb{N}}$ est un **processus** qui admet \mathbb{N} (ou un intervalle de \mathbb{N}) comme ensemble des **temps**.

L'ensemble E des **états** possibles de ce processus est contenu dans \mathbb{Z} : c'est un ensemble fini ou dénombrable. Pour chaque valeur k du temps, X_k "dépend du hasard" : c'est une **variable aléatoire**. X_k n'a pas de valeur précise : par contre on peut étudier la **probabilité** pour que X_k prenne la valeur j ce qu'on note $\text{Proba}[X_k = j]$.

b) Pierre marche sur une route en ligne droite parallèle à l'axe Nord-Sud ; tous les kilomètres il joue à pile ou face ; si la pièce retombe sur pile, il fait un kilomètre vers le Nord ; si la pièce retombe sur face, il fait un

kilomètre vers le Sud. Soit X_k l'abscisse de Pierre, en kilomètres, au k -ième kilomètre parcouru, l'origine étant la position initiale de Pierre et la route étant orientée du Sud au Nord. La loi d'évolution de X_k est évidemment la même qu'au a). On dit que X est une **marche aléatoire**.

c) Si on connaît X_k , la valeur de X_{k+1} ne dépend pas des valeurs $X_1, \dots, X_j, \dots, X_{k-1}$: on dit que le processus est **markovien** (cf. § B.3).

d) Soit Y_k le nombre de fois où la pièce est tombée sur pile lors des k premiers lancers. On sait que Y_k suit la **loi binomiale** d'ordre k et de paramètre $1/2$. On a donc, pour $0 \leq j \leq k$,

$$\text{Proba}[Y_k = j] = C_k^j \left(\frac{1}{2}\right)^k \quad \text{où} \quad C_n^j = \frac{n!}{j!(n-j)!}$$

On a évidemment $\text{Proba}[X_k = i] = \text{Proba}[Y_k = (i+k)/2]$.

A.3 Loi et processus de Poisson

a) On étudie les appels qui arrivent à un standard téléphonique. On effectue cette étude pendant une heure. On suppose que, durant chaque intervalle de temps de une seconde, la probabilité qu'il y ait un nouvel appel vaut p , et la probabilité qu'il n'y ait aucun appel vaut $(1-p)$: on suppose donc que la probabilité qu'il y ait plusieurs appels est négligeable. On suppose aussi que l'évolution est **markovienne** c'est à dire que ce qui se passe pendant une seconde est **indépendant** de ce qui s'est passé **auparavant**. Soit Y le nombre d'appels arrivés au cours de l'heure considérée. On sait que Y suit la **loi binomiale** d'ordre 3.600 et de paramètre p . La moyenne (l'espérance mathématique) de Y vaut $a = p n$ avec $n := 3.600$.

$$\text{On a} \quad \text{proba}[Y = k] = C_n^k p^k (1-p)^{n-k}$$

n étant grand, l'expression ci-dessus est techniquement difficile à calculer. Dans ce cas, on sait que la loi binomiale peut être approchée par la **loi de Poisson** de même moyenne a . Rappelons que l'on dit que Z suit la loi de poisson de paramètre a , si on a :

$$\text{proba}[Z = k] = e^{-a} a^k / k!$$

b) Etudions maintenant le même central mais en observant les arrivées d'appels milliseconde par milliseconde. L'étude peut être effectuée comme au a) sauf que n doit être remplacée par $n' = n 10^3$ et p

par $p' = p \cdot 10^{-3}$ puisque la moyenne $a = np = n'p'$ n'a pas changée. Dans ce cas, l'approximation par la loi de Poisson est encore meilleure qu'au a).

On conçoit donc qu'en observant les nouveaux appels suivant des intervalles de temps de plus en plus brefs, on ait, à la limite, **exactement une loi de Poisson**.

c) Nous verrons (§ E.3) une autre façon de démontrer ce résultat. Au lieu d'étudier le nombre d'appels durant une heure, on peut l'étudier entre 0 et t .

Plus précisément, soit $T = \mathbb{R}^+$ l'ensemble des temps ; pour chaque élément t de T , soit X_t la variable aléatoire associée aux nombres d'appels arrivés entre 0 et t . Le processus $(X_t)_{t \in T}$ est appelé **processus de Poisson**.

A.4 Loi exponentielle

a) Reprenons le "décor" considéré au paragraphe A.3 mais, au lieu d'étudier le nombre d'appels au cours d'une heure, nous allons étudier le premier appel. La méthode la plus classique consiste à étudier la variable aléatoire D associée au délai qui s'écoule avant le premier appel ; on sait que D suit la **loi exponentielle** de paramètre u ; le délai entre deux appels suit la même loi que D ; si D est exprimé en heures, $u = 1/a$ (a étant le paramètre introduit en A.3).

On peut aussi aborder ce problème en termes de processus. Plus précisément, 0 étant l'instant initial d'observation, pour chaque instant t on pose $X_t = 0$ s'il n'y a pas eu d'appel avant l'instant t et $X_t = 1$ si le premier appel a eu lieu avant l'instant t . L'ensemble des temps de ce processus X est \mathbb{R}^+ ; l'**ensemble des états** de ce processus X est $E := \{0, 1\}$. Quel que soit $t \geq 0$, X_t est une **variable aléatoire** et on a :

$$\text{Proba}[X_t = 0] = \text{Proba}[D > t] = e^{-ut} \quad \text{et}$$

$$\text{Proba}[X_t = 1] = \text{Proba}[D \leq t] = 1 - e^{-ut}$$

On a aussi :

$$\text{Proba}[X_{t+s} = 0 \mid X_s = 0] = e^{-u(t+s)} / e^{-us} = e^{-ut} = \text{Proba}[X_t = 0]$$

c'est à dire que, à partir de l'instant s , si on sait que $X_s = 0$, l'évolution ultérieure (s fixe, t variable) suit la même loi que celle du processus X_t initial : cette propriété est satisfaite par tous les **processus markoviens homogènes**, c'est à dire les processus markoviens dont les taux d'évolution (u pour X) sont fixes au cours du temps. L'application $X_t \rightarrow X_{t+s}$ est

souvent appelée **opérateur de translation**.

b) Considérons une machine que l'on met en marche à l'instant $t = 0$; on veut étudier le premier instant où cette machine tombe en panne. On pose $X_t = 0$ si la machine n'est pas tombée en panne avant l'instant t et $X_t = 1$ dans le cas contraire. On suppose que, si la machine est en marche à l'instant s , la probabilité pour qu'elle tombe en panne entre s et $s+h$ vaut $u h + h \varepsilon(h)$ où $\varepsilon(h)$ est une fonction de h telle que

$$\lim_{h \downarrow 0} \varepsilon(h) = 0 .$$

On suppose donc que u ne varie pas, c'est à dire que la machine considérée ne "**vieillit pas**". On sait alors que X_t suit la même loi qu'au a) (étude de la loi exponentielle).

Si la machine étudiée est une voiture, au lieu d'étudier le délai avant la première panne, on peut étudier le kilométrage parcouru avant la première panne (t sera exprimé en kilomètres).

Comme en A.2, on note qu'un même modèle mathématique peut convenir pour des situations concrètes extrêmement diverses.

B. Loi d'un processus

B.1 Processus stochastique

En théorie des probabilités à temps fixe, l'expression peut-être la plus utilisée est celle de "**variable aléatoire**" : en fait, il n'y a pas de définition mathématique simple et universelle de cette expression. Cette expression correspond à un concept, à la frontière entre la modélisation mathématique et le réel ; par contre, si R et S sont deux variables aléatoires, on peut donner un sens mathématique très précis à des notations ou expressions telles que :

Proba [$S = u$] qui peut se lire "probabilité d'avoir $S = u$ "

$E(S)$ = espérance mathématique (valeur moyenne) de S

Proba [$S = u | R = v$] qui peut se lire "probabilité d'avoir $S = u$ sachant $R = v$ ", etc...

De même, si on étudie un phénomène qui évolue au cours du temps (ou au fil des kilomètres !), le concept de base est celui de **processus**. Evidemment, on peut donner des définitions mathématiques précises associées à ce concept : ces définitions dépendent de ce que l'on veut

étudier et du niveau théorique auquel on se situe. En fait, de même que pour l'expression "variable aléatoire", l'expression "**processus stochastique**" correspond à un concept à la frontière entre les mathématiques et le domaine expérimental.

Nous ne donnerons pas de définition mathématique de l'expression "processus stochastique" (en abrégé processus). Disons seulement que, si T est l'ensemble des temps, un processus (stochastique) $(X_t)_{t \in T}$ est une famille de variables aléatoires indexée par T . Pour tout élément t de T , X_t est la variable aléatoire associée à l'état du processus à l'instant t . Notamment $\text{Proba}[X_t = u]$ désigne la probabilité qu'à l'instant t le processus soit dans l'état u . De même, pour $s < t$, $\text{Proba}[X_t = u | X_s = v]$ désigne la probabilité que le processus X soit dans l'état u à l'instant t sachant qu'il était dans l'état v à l'instant s .

Dans ce cours, l'ensemble T des temps sera toujours un intervalle de \mathbb{N} (cas "à temps discret") - exemple A.2 - ou de \mathbb{R} (cas "à temps continu") - exemples A.3.c) et A.4. L'ensemble E des états du processus sera toujours un ensemble fini ou dénombrable (processus "discret").

B2 Loi d'un processus

Soit $t(1) < t(2) < \dots < t(n)$ une séquence d'éléments de T ; la famille $(X_{t(1)}, X_{t(2)}, \dots, X_{t(n)})$ peut être considérée comme une variable aléatoire à valeurs dans E^n et on peut parler de sa **loi conjointe**.

Etudier la **loi du processus** X , c'est étudier la famille de toutes les lois conjointes des variables $X_{t(1)}, \dots, X_{t(n)}$ pour toutes les séquences finies croissantes possibles d'éléments de T . Cette famille satisfait évidemment à des conditions naturelles de compatibilité (par exemplar la loi de (X_u, X_v, X_w) induit celle de (X_u, X_v)) : on dit qu'elle constitue un **système projectif**.

Il est très vite apparu que cette notion de loi ne suffisait pas, au niveau de la technique mathématique, quand l'ensemble T des temps est un intervalle de la droite réelle : notamment elle ne suffit pas pour disposer de l'outil indispensable que le mathématicien appelle un temps d'arrêt. Le formalisme adéquat est dû essentiellement à Kolmogorov (cf. le chapitre 10 et le paragraphe 3.E.4).

B.3 Processus markovien

On dit qu'un processus $(X_t)_{t \in T}$ est **markovien** si "**le passé et le futur sont indépendants sachant le présent**", c'est à dire que l'évolution (la loi) du processus à partir de l'instant s , sachant qu'il est dans l'état e à cet instant, ne dépend pas de la façon dont ce processus a atteint cet état e mais seulement de l'état e .

Dans le cas où l'ensemble des états est fini ou dénombrable - ce qui est "presque" le seul cas considéré dans ce cours - soit $t(1) < t(2) < \dots < t(m) < \dots < t(n)$ une séquence d'éléments de T et $(e_1, \dots, e_m, \dots, e_n)$ une séquence associée d'éléments de E ; dire que le processus X est markovien, c'est dire que pour toutes telles séquences on a :

$$\begin{aligned} \text{Proba} [\forall i > m, X_{t(i)} = e_i \mid X_{t(m)} = e_m] \\ = \text{Proba} [\forall i > m, X_{t(i)} = e_i \mid \forall j \leq m, X_{t(j)} = e_j] \end{aligned}$$

B.4 Semi-groupe de transition

Si la propriété ci-dessus est satisfaite, on vérifie immédiatement en raisonnant par récurrence sur m et en utilisant la formule de Bayes que

$$\text{Proba} [\forall i, X_{t(i)} = e_i] = \text{Proba} [X_{t(1)} = e_1] \cdot \prod_{k=2}^m \text{Proba} [X_{t(k)} = e_k \mid X_{t(k-1)} = e_{k-1}]$$

On a de plus (σ -additivité d'une probabilité qu'on appelle aussi axiome des probabilités totales) :

$$\text{Proba} [X_{t(1)} = e_1] = \sum_{e \in E} \text{Proba} [X_{t(0)} = e] \cdot \text{Proba} [X_{t(1)} = e_1 \mid X_{t(0)} = e]$$

La fonction $G_{s,t}(e,e') := \text{Proba} [X_t = e' \mid X_s = e]$, définie pour $s \leq t$, est appelée **semi-groupe de transition**. Ce qui précède montre que la loi d'un processus markovien à partir de l'instant $t(0)$ est caractérisée par la **loi initiale** (la loi de la variable $X_{t(0)}$) et le semi-groupe de transition G .

On dit que le processus X markovien est **homogène** si, quel que soit $u > 0$ et $s < t$, $G_{s,t} = G_{s+u,t+u}$ (dans la mesure où s et $t+u$ appartiennent à T). La loi du processus est invariante par translation dans le temps.

Dans la suite, du point de vue des applications, on se ramènera toujours à l'étude de **processus markoviens homogènes discrets**. Notamment, si on étudie un tel processus par simulation informatique

(cf. § 3.D), l'ensemble des états possibles est grand mais fini. Par contre, on est souvent amené à introduire un modèle mathématique pour lequel l'ensemble des états est non dénombrable : par exemple, dans le cas de la file GI/GI/1 (cf. 3.E.3), on met en mémoire les délais passés depuis la dernière arrivée et le début du dernier service.

C. Chaîne de Markov

C.1 Matrice de transition

On dit qu'un processus $(X_t)_{t \in T}$ est une **chaîne de Markov** si $(X_t)_{t \in T}$ est un processus de Markov et si l'ensemble T des temps est un intervalle de \mathbb{N} . Dans ce cas, la valeur du semi-groupe de transition pour $s = k$ et $t = k+1$ est appelée "**matrice de transition**".

Plus précisément, si E est fini, on peut identifier E et l'ensemble $\{1, \dots, n\}$. Pour tout élément k de $T \subset \mathbb{N}$, la matrice de transition en k est la matrice $(n \times n)$ définie par :

$$(g_k)_{i,j} := \text{Proba} [X_{k+1} = j \mid X_k = i] .$$

Le processus est homogène si cette matrice ne dépend pas de k .

C.2 Semi-groupe de transition

Soit $(X_t)_{t \in T}$ une chaîne de Markov dont l'ensemble des états est $\{1, \dots, n\}$. Soit $h < k$ deux éléments de T . On a :

$$\text{Proba} [X_k = j \mid X_h = i] = (g_h g_{h+1} \cdots g_{k-1})_{i,j}$$

l'expression ci-dessus entre parenthèses étant un produit matriciel usuel.

Ceci se vérifie immédiatement (comme en B.4) en raisonnant par récurrence sur k et en utilisant l'additivité d'une probabilité et la formule de Bayes.

C.3 Matrice stochastique

Soit g une matrice $(n \times n)$ à coefficients positifs. On dit que g est une **matrice stochastique** si, quel que soit i , $1 \leq i \leq n$, on a :

$$\sum_{j=1}^n g_{i,j} = 1$$

On vérifie immédiatement que le produit de deux matrices stochastiques est une matrice stochastique et qu'une matrice de transition (comme définie en C.1) est une matrice stochastique. Inversement, une matrice stochastique peut être considérée comme une matrice de transition.

Nous allons donner trois exemples de chaînes de Markov ; de nombreux ouvrages proposent d'autres exemples tout aussi illustratifs (cf., par exemple, [Fau], [HeP] et [Ros] ; cf. aussi 3.D.3).

Le premier exemple proposé en C.4 est tout à fait élémentaire. En C.5 on donne un exemple de stratégie ; en C.6 on donne un exemple de gestion de stock : ce sont des situations types pour lesquelles la modélisation mathématique associée est une chaîne de Markov.

C.4 Pile ou Face

Jean et Pierre jouent à Pile ou face. Au début de la partie, Jean possède a francs et Pierre b francs. A chaque partie, Jean et Pierre misent un franc chacun et celui des deux qui gagne ramasse la mise totale. Si X_k est la fortune de Jean après le k -ième partie, celle de Pierre est $b+a-X_k$ donc la valeur de X_k caractérise l'état du processus. Tant que $0 < X_k < a+b$, on a $\text{Proba}[X_{k+1} = u+1 \mid X_k = u] = \text{Proba}[X_{k+1} = u-1 \mid X_k = u] = 1/2$, ce qui définit la matrice de transition. Il est assez facile d'étudier complètement la loi de ce processus.

C.5 Une stratégie élémentaire

Soit $b > 0$ et m et n deux entiers, $m > n > 1$. Soit $(X_k, Y_k)_{k \geq 0}$ la suite de couples de variables aléatoires définie comme suit :

- si $b < X_k < b \cdot 2^n$, $Y_{k+1} = Y_k$
et $\text{Proba}[X_{k+1} = 2X_k] = \frac{1}{3}$ et $\text{Proba}[X_{k+1} = \frac{1}{2} X_k] = 2/3$
- si $X_k = b$, $\text{Proba}[X_{k+1} = b \text{ et } Y_k = Y_k - \frac{b}{2}] = 2/3$
et $\text{Proba}[X_{k+1} = 2b \text{ et } Y_k = Y_k] = 1/3$
- si $X_k = b \cdot 2^n$, $\text{Proba}[X_{k+1} = \frac{1}{2} X_k \text{ et } Y_{k+1} = Y_k] = 2/3$
et $\text{Proba}[X_{k+1} = X_k \text{ et } Y_{k+1} = b \cdot 2^n + Y_k] = 1/3$

Ceci modélise l'état de la fortune d'un joueur qui adopte la stratégie ("martingale") suivante : soit X_k la somme destinée au jeu et Y_k le reste de la fortune juste avant la k -ième partie. Lors de cette k -ième partie, si $b < X_k < 2^n b$, il joue la somme $\frac{1}{2} X_k$ et on suppose qu'il a deux chances sur trois pour que la somme destinée au jeu passe de X_k à $\frac{1}{2} X_k$ (s'il perd) et une chance sur trois pour que cette somme passe de X_k à $2X_k$ (s'il gagne) : si $b < X_k < 2^n b$, il ne modifie pas le reste de sa fortune Y_k . Si $X_k = b$, il joue encore la somme $\frac{1}{2} X_k$ mais, s'il perd, il prend la somme $\frac{1}{2} X_k = \frac{1}{2} b$ dans sa fortune et la rajoute à la somme $\frac{1}{2} b$ qui lui servait pour jouer. Si $X_k = b.2^n$, il joue toujours la somme $\frac{1}{2} X_k$ mais, s'il gagne, il remet prudemment la somme $b.2^n$ dans sa fortune et ne garde que la somme $b.2^n$ pour jouer le coup suivant.

On suppose que $Y_1 = b.2^n$ et $X_1 = b$. Enfin, le joueur décide d'arrêter de jouer dès que $Y_k = 0$ ou $Y_k \geq b.2^{m+1}$.

La valeur du couple (X_{k+1}, Y_{k+1}) dépend du hasard à la k -ième partie et de la valeur du couple (X_k, Y_k) ; par contre, **si on connaît ce couple**

(X_k, Y_k) , (X_{k+1}, Y_{k+1}) **ne dépend pas des couples (X_j, Y_j) pour $j < k$** . On a donc une **chaîne de Markov** si on prend comme ensemble d'états l'ensemble des valeurs possibles de ces couples (X_k, Y_k) . Si on voulait définir la "matrice" de transition avec le formalisme usuel, il faudrait donner une indexation totalement ordonnée de cet ensemble d'états. En fait, dans l'essentiel de ce cours, la relation d'ordre sur les indices d'une matrice n'intervient presque jamais (cf. le chapitre 5 pour une explicitation du formalisme général associé). On va donc considérer qu'un état u est un couple $u := (u', u'')$ où u' est de la forme $2^k b$, $0 \leq k \leq n$ et u'' est de la forme $j b$, $0 \leq j \leq 2^{m+1}$. Pour $b < u' < 2^n b$, on a :

$$\text{Proba} [(X_{k+1}, Y_{k+1}) = (u'/2, u'') | (X_k, Y_k) = (u', u'')] = 2/3$$

$$\text{Proba} [(X_{k+1}, Y_{k+1}) = (2u', u'') | (X_k, Y_k) = (u', u'')] = 1/3$$

Les autres probabilités de transition pour $b < u' < 2^n b$ sont nulles. On construit, de façon analogue, $a_{u,v}$ pour $u' = b$ et $u' = 2^n b$.

On note que le jeu est équilibré en ce sens que $E(X_{k+1} + Y_{k+1} - X_k - Y_k) = 0$,

et ceci quelle que soit la valeur du couple (X_k, Y_k) : le mathématicien dit alors que $(X_k + Y_k)$ est une **martingale**.

En fait, nous verrons plus tard que l'on peut être amené à "grossir" l'ensemble des états si on veut disposer **d'autres informations que la fortune du joueur** : par exemple, si on veut étudier le **nombre total de parties** avant que le joueur n'arrête de jouer, il faut que **l'état du processus prenne en compte** le nombre de parties déjà jouées. Cet exemple sera repris en 2.C.2, 2°) et surtout, en 3.D.2.

C.6 Stock à point de commande

Soit u et v deux entiers strictement positifs.

Tous les lundis matin, un détaillant fait l'inventaire de son stock : pour simplifier, supposons qu'il n'y ait qu'un seul article. Si le stock X_k est supérieur ou égal à v , le détaillant ne commande rien. Si le stock est inférieur à v , le détaillant commande j articles où j est le plus petit entier i tel que $X_k + iu \geq v$. Pour simplifier le formalisme, on suppose que ces articles sont livrés immédiatement (le lundi matin).

Par ailleurs, on suppose qu'une étude statistique préliminaire a montré que le nombre Y_k d'articles vendus (s'ils sont disponibles) entre deux lundis matin suit une loi de Poisson de paramètre a .

Pour tout entier w , posons :

$$a(w) := \inf. \{i : w + iu \geq v\} \text{ et } b(w) := w + a(w)$$

c'est à dire que le détaillant commande $a(w)$ articles si le stock est dans l'état w (le lundi matin).

La matrice de transition est alors définie par :

$$\text{Proba}[X_{k+1} = 0 \mid X_k = w] = e^{-a} \sum_{j \geq b(w)} a^j / j !$$

et pour $b(w) - j \geq 1$

$$\text{Proba}[X_{k+1} = b(w) - j \mid X_k = w] = e^{-a} a^j / j !$$

Pour un tel exemple, le problème essentiel est **d'optimiser, en régime stationnaire** (cf. chapitre 2) les valeurs de u et v relativement à une certaine **fonction de coût**.

D. Processus markovien à temps continu

D.1 Matrice d'évolution

Dans tout ce paragraphe D, on pose $T := \mathbb{R}^+$, on suppose que E est un **ensemble fini ou dénombrable** et on suppose que $(X_t)_{t \in T}$ est un processus markovien homogène qui admet E comme ensemble d'états. On pose :

$$(1D1) \quad g_t(u, v) := \text{Proba} [X_t = v \mid X_0 = u] = \text{Proba} [X_{t+s} = v \mid X_s = u]$$

On suppose que, pour tout couple (u, v) d'éléments de E , la fonction $h \rightarrow \frac{1}{h} g_h(u, v)$ a une limite quand h tend vers zéro, $h > 0$: c'est une hypothèse de "dérivabilité à droite" physiquement assez naturelle. On pose, pour $u \neq v$:

$$(1D2) \quad a(u, v) := \lim_{h \downarrow 0} \frac{1}{h} g_h(u, v) = \lim_{h \downarrow 0} \frac{1}{h} \text{Proba} [X_{t+h} = v \mid X_t = u]$$

Pour $u = v$, on pose $a(u, u) := 0$.

On dira que a est la **matrice d'évolution** associée au processus X .

D.2 Equations de Chapman-Kolmogorov

On suppose que l'ensemble E est fini.

Soit $h > 0$. Provisoirement, on fixe t , u , v et h et, pour alléger les notations, on pose :

$$z(w) := \text{Proba} [X_{t+h} = v \text{ et } X_t = w \mid X_0 = u]$$

Le processus X étant markovien, on a :

$$z(w) = g_t(u, w) g_h(w, v) \quad (\text{formule de Bayes})$$

Pour $w = v$, on a aussi :

$$z(v) = g_t(u, v) [1 - \sum_{w \neq v} g_h(v, w)] \quad \text{puisque} \quad \sum_{w \in E} g_h(v, w) = 1$$

Enfin, l'additivité de toute loi de probabilité implique :

$$g_{t+h}(u, v) = \sum_{w \in E} z(w) = \sum_{w \neq v} g_t(u, w) g_h(w, v) + g_t(u, v) [1 - \sum_{w \neq v} g_h(v, w)]$$

ce qui permet aussi d'écrire :

$$\frac{1}{h} [g_{t+h}(u,v) - g_t(u,v)] = - g_t(u,v) \sum_{w \neq v} \frac{1}{h} g_h(v,w) + \sum_{w \neq v} g_t(u,w) \frac{1}{h} g_h(w,v)$$

Compte tenu de D.1 (cf. 1D2), le second membre a une limite quand h tend vers zéro ; le premier membre a donc lui aussi une limite, c'est à dire que la fonction $t \rightarrow g_t(u,v)$ est dérivable et on a :

$$(1D3) \quad g'_t(u,v) = -g_t(u,v) \sum_{w \neq v} a(v,w) + \sum_{w \neq v} g_t(u,w) a(w,v)$$

Ce système différentiel est souvent appelé le système des équations de Chapman-Kolmogorov. Il reste valable quand E est infini sous réserve d'être autorisé à intervertir le symbole Σ et le passage à la limite quand h tend vers zéro. En général, quand on a à écrire ce système d'équations pour un exemple concret, il ne faut pas chercher à utiliser formellement la relation 1D3 : au contraire, il est souvent plus simple de refaire la démonstration qui précède pour l'exemple considéré. On conseille donc au lecteur novice de **refaire cette démonstration** pour tous les exemples considérés par la suite, notamment ceux du paragraphe E (cf. aussi les exemples du chapitre 4, 5.E.5, 9.C.2, etc...)

D.3 Equations backward de Kolmogorov

Les équations 1D3 sont aussi appelées les **équations forward de Kolmogorov**. En introduisant l'évolution entre 0 et h au lieu de l'évolution entre t et $t+h$ on démontre, comme en D.2, que l'on a :

$$\begin{aligned} g_{t+h}(u,v) &= \sum_{w \in E} \text{Proba}[X_{t+h} = v \text{ et } X_h = w \mid X_0 = u] \\ &= \sum_{w \neq u} g_h(u,w) g_t(w,v) + g_t(u,v) [1 - \sum_{w \neq u} g_h(u,w)] \end{aligned}$$

En divisant par h et en faisant tendre h vers zéro, on obtient :

$$(1D4) \quad \begin{aligned} g'_t(u,v) &= -g_t(u,v) \sum_{w \neq u} a(u,w) \\ &\quad + \sum_{w \neq u} a(u,w) g_t(w,v) \end{aligned}$$

ces équations sont souvent appelées les **équations backward de Kolmogorov**.

D.4 Système différentiel linéaire

Théorème : E est un ensemble fini ou dénombrable. On se donne une constante c et une fonction réelle b définie sur $(E \times E)$. Suivant les

conventions générales de ce cours, on pose $b_{u,v} := b(u,v)$.

On suppose que, pour tout élément u de E , on a :

$$(1D5) \quad \sum_{v \in E} |b_{u,v}| \leq c$$

Alors il existe une et une seule famille de fonctions réelles de variables réelles, indéfiniment dérivables, famille indexée par $(E \times E)$ que l'on note (e^{bt}) et qui satisfait aux propriétés suivantes :

- (i) pour tout couple (u,v) d'éléments de E , les familles $(b_{u,w}(e^{bt})_{w,v})_{w \in E}$ et $((e^{bt})_{u,w} b_{w,v})_{w \in E}$ sont sommables et on a :

$$\frac{d}{dt} (e^{bt})_{u,v} = \sum_{w \in E} b_{u,w} (e^{bt})_{w,v} = \sum_{w \in E} (e^{bt})_{u,w} b_{w,v}$$

- (ii) $(e^0)_{u,v} = \delta_{u,v}$ (symbole de Kronecker)

- (iii) pour tout couple (u,v) d'éléments de E et pour tout couple (s,t) de réels la famille $((e^{bs})_{u,w} (e^{bt})_{w,v})_{w \in E}$ est sommable et

$$(e^{b(s+t)})_{u,v} = \sum_{w \in E} (e^{bs})_{u,w} (e^{bt})_{w,v}$$

Preuve :

1°) Ce théorème est un résultat classique d'analyse que l'on peut trouver dans tous les ouvrages de base quand E est fini. Pour la commodité du lecteur, rappelons rapidement les grandes étapes de la preuve (quand E est infini) : cette preuve peut être omise en première lecture (cf. aussi 8.A.2).

2°) Pour tout couple (F,G) de parties finies de E , on a :

$$\sum_{v \in F} \sum_{w \in G} |b_{u,v} b_{v,w}| \leq c \sum_{v \in F} |b_{u,v}| \leq c^2$$

La famille $(b_{u,v} b_{v,w})_{v \in E}$ est donc sommable et on peut poser

$$b_{u,w}^2 := \sum_{v \in E} b_{u,v} b_{v,w} \neq (b_{u,v})^2$$

3°) En raisonnant par récurrence croissante sur k , on prouve de même que l'on peut poser :

$$b_{u,w}^k := \sum_{v \in E} b_{u,v} (b^{k-1})_{v,w} \quad \text{et on a}$$

$$b_{u,w}^k = \sum_{v \in E} (b^{k-1})_{u,w} b_{w,v} \quad \text{et}$$

$$\sum_{w \in E} |b_{u,w}^k| \leq c^k$$

4°) La série entière de terme général $b_{u,v}^k t^k/k!$ est majorée en module par $(|t| c)^k/k!$; ces deux séries ont un rayon de convergence infini. Notamment, on peut poser

$$(e^{bt})_{u,v} := \sum_{k=0}^{\infty} b_{u,v}^k t^k/k!$$

5°) En raisonnant comme précédemment, on vérifie que les familles considérées en (i) sont sommables. On termine la preuve de (i) en dérivant termes à termes la série entière introduite au 4°).

6°) La propriété (ii) est évidente. On vérifie que la famille introduite en (iii) est sommable. Par dérivation on prouve (iii) quand $s = -t$. On en déduit l'unicité de la famille e^{bt} en tant que solution de (i) et (ii).

7°) On prouve (iii) par dérivation et identification (unicité) (pour plus de détails, cf. 8.A).

D.5 Semi-groupe de transition

1°) E est un ensemble fini ou dénombrable et on pose $T := \mathbb{R}^+$. Soit c' une constante et a une fonction positive définie sur $(E \times E)$. On suppose que, pour tout élément u de E , on a :

$$(1D6) \quad \sum_{v \in E} a(u,v) \leq c'.$$

On pose $c := 2c'$. Soit b la fonction définie sur $(E \times E)$ par :

$$\text{si } u \neq v, \quad b(u,v) = a(u,v)$$

$$\text{si } u = v, \quad b(u,u) = - \sum_{v \in E} a(u,v)$$

On note que $\sum_{v \in E} |b(u,v)| \leq c$. On peut donc utiliser le théorème D.4. On pose

$$g_t(u,v) := (e^{bt})_{u,v}.$$

2°) La famille de fonctions $g_t(u,v)$ est l'unique famille qui satisfait aux équations 1D3 (resp. 1D4) et aux conditions initiales $g_0(u,v) = \delta_{u,v}$.

On dit souvent que cette famille est le semi-groupe de transition : en fait, dans le cas considéré ici, cette famille constitue un groupe quand t parcourt l'axe réel, le cas $t < 0$ n'ayant plus d'interprétation "physique" naturelle. On peut même prendre t complexe.

3°) Si on connaît la loi initiale $(p_u)_{u \in E}$ avec $p_u := \text{Proba}[X_0 = u]$, on a :

$$\text{Proba}[X_t = v] = \sum_{u \in E} p_u (e^{bt})_{u,v}$$

On vérifie immédiatement que la fonction $f(t) := \sum_{v \in E} \text{Proba}[X_t = v]$ a une dérivée nulle ; elle est donc constante et égale à 1.

4°) L'hypothèse $\sum_{v \in E} a(u,v) \leq c'$ avec c' indépendant de u est une hypothèse forte qui n'est pas toujours satisfaite (cas de la file $M/M/\infty$ par exemple). Quand cette hypothèse n'est pas satisfaite, en général, le plus simple est de considérer une suite de fonctions $(a_k)_{k>0}$ qui croît vers a et telle que chaque fonction a_k satisfasse à l'hypothèse précitée. On vérifie alors, quand c'est possible, que la suite des solutions associées à $(a_k)_{k>0}$ converge vers la solution escomptée quand k tend vers l'infini.

E. Exemples

E.1 Loi exponentielle

Reprenons l'exemple introduit en A.4. On a $E = \{0,1\}$, $a_{0,1} = u$, $a_{1,0} = 0$. Le système différentiel 1D3 s'écrit, pour $u = 0$ et $u = 1$:

$$\begin{aligned} \frac{d}{dt} g_t(u,0) &= -u g_t(u,0) \\ \frac{d}{dt} g_t(u,1) &= u g_t(u,0) \end{aligned}$$

Pour $u = 0$, $g_t(0,0) = 1 - e^{-ut}$ et $g_t(0,1) = e^{-ut}$.

Pour $u = 1$, $g_t(1,0) = 0$ et $g_t(1,1) = 1$.

E.2 Loi d'Erlang

1°) Soit $u > 0$. Considérons le cas où $E := \{0, 1, \dots, n\}$. On suppose que, quel que soit k , $0 \leq k \leq n-1$, $a(k, k+1) := u$ et que $a(e, e') := 0$ pour tous les autres couples (e, e') d'éléments de E . Ceci généralise le cas précédent qui correspond à $n = 1$.

Ce modèle correspond, par exemple, au cas d'un client qui passe successivement dans n services, la durée de chaque service suivant la loi exponentielle de paramètre u . Pour $k < n$, le processus X est dans l'état k si le client est dans le $(k+1)$ -ième service. Le processus est dans l'état n si les n services sont terminés.

2°) Comme précédemment, on pose :

$$g_t(j, k) := \text{Proba}[X_t = k \mid X_0 = j]$$

Le système 1D3 s'écrit :

$$g'_t(j, 0) = -u g_t(j, 0), \quad g'_t(j, n) = u g_t(j, n-1)$$

et pour $0 < k < n$:

$$g'_t(j, k) = -u g_t(j, k) + u g_t(j, k-1)$$

3°) Pour $j = 0$, on a $g_0(0, k) = \delta_{0,k}$. En raisonnant par récurrence croissante sur k on vérifie facilement que la solution de ce système différentiel est :

$$\text{pour } 0 \leq k < n : g_t(0, k) = (ut)^k e^{-ut} / k!$$

$$\text{et } g_t(0, n) = 1 - \sum_{k=0}^{n-1} (ut)^k e^{-ut} / k!$$

$g_t(0, n)$ est la **fonction de répartition de la loi d'Erlang d'ordre n** , c'est à dire du produit de convolution de n lois exponentielles.

4°) Pour $j > 0$ et $j \leq k \leq n$, on a $g_t(j, k) = h_t(0, k-j)$ où h_t est la fonction g_t sauf qu'on y remplace n par $(n-j)$: on a une loi d'Erlang $(n-j)$ puisque les j premiers services sont déjà effectués dès l'instant $t = 0$. Evidemment, $g_t(j, k) = 0$ pour $k < j$.

E.3 Processus de Poisson

Reprenons le "décor" considéré en A.3 avec la modélisation amorcée en A.3.c. On a donc $X_t = k$ s'il y a eu k appels entre 0 et t ; $E = \mathbb{N}$ et $T = \mathbb{R}^+$. On a $a(k, k+1) = u$ (quel que soit $k \geq 0$) et $a(e, e') = 0$ pour tous les autres

couples d'éléments de E . L'étude est tout à fait analogue à celle effectuée en E.2 . Le système différentiel 1D3 s'écrit :

$$g'_t(j,0) = -u g_t(j,0) \quad \text{et, pour } k > 0 :$$

$$g'_t(j,k) = -u g_t(j,k) + u g_t(j,k-1)$$

On en déduit :

$$\text{pour } k < j, \quad g_t(j,k) = 0$$

$$\text{pour } k \geq j, \quad g_t(j,k) = (ut)^{(k-j)} e^{-ut} / (k-j)!$$

E.4 File M/M/1

1°) Définissons d'abord le modèle mathématique. On pose $T := \mathbb{R}^+$ et $E := \mathbb{N}$. Soit $u > 0$ et $v > 0$. On pose, quel que soit $k \geq 0$:

$a(k,k+1) := u$, $a(k+1,k) := v$ et $a(e,e') := 0$ pour tous les autres couples (e,e') d'éléments de E . Le système 1D3 donne :

$$g'_t(j,0) = -u g_t(j,0) + v g_t(j,1)$$

et, pour $k > 0$:

$$g'_t(j,k) = -(u+v) g_t(j,k) + u g_t(j,k-1) + v g_t(j,k+1)$$

On sait calculer la solution de ce système différentiel mais celle-ci est loin d'être simple : on note donc que, même **pour un système différentiel 1D3 relativement simple, la solution est complexe et souvent peu utilisable.**

2°) Ce modèle mathématique correspond à des situations concrètes extrêmement variées. Par exemple, k peut être le nombre de clients dans une file d'attente (y compris le client en service) quand les arrivées à cette file se font suivant un processus de Poisson de paramètre u et quand il y a un seul serveur dont la durée de service suit la loi exponentielle de paramètre v . Dans la **notation M/M/1**, le premier M correspond à la loi des arrivées (Markov), le deuxième M à la loi de service (Markov) et le 1 au fait qu'il y a un seul serveur.

E.5 File E2/M/1

1°) Définissons d'abord le modèle mathématique. E , l'ensemble des états, est l'ensemble des couples (j,k) avec $j = 1$ ou $j = 2$ et k entier positif. Soit u et v deux réels strictement positifs. On pose, quel que soit $k \geq 0$:

$$a((1,k),(2,k)) := u, \quad a((2,k),(1,k+1)) := u$$

$$a((1,k+1),(1,k)) := v, \quad a((2,k+1),(2,k)) := v$$

Pour tous les autres couples d'états (e, e') on pose $a(e, e') := 0$. On pose aussi $g(e, (j, -1)) := 0$ (quel que soit l'état e et $j = 1$ ou $j = 2$). Le système 1D3 devient dans ce cas, quels que soient l'état e et $k > 0$:

$$g'(e, (1, k)) = -g(e, (1, k))(u+v) + g(e, (2, k-1)) u + g(e, (1, k+1)) v$$

$$\text{et} \quad g'(e, (2, k)) = -g(e, (2, k))(u+v) + g(e, (1, k)) u + g(e, (2, k+1)) v$$

Si $k = 0$, dans les relations ci-dessus, on remplace $(u+v)$ par u .

La solution explicite de ce système est compliquée et peu utilisable.

2°) Ce modèle mathématique correspond, par exemple, à la situation "physique" suivante : un automate oscille constamment entre les positions 1 et 2, ce qui correspond aux valeurs $j = 1$ et $j = 2$. Le temps passé dans chacun de ces états suit la loi exponentielle de paramètre u . A chaque fois que l'automate passe de la position 2 à la position 1, il provoque l'arrivée d'un client dans une station S (il ouvre la porte de S et laisse rentrer un client) ; cette station S comporte un seul serveur dont le délai de service suit la loi exponentielle de paramètre v (comme en E.4) : k correspond au nombre de clients dans la file.

3°) Le modèle mathématique considéré au 1°) peut aussi servir à modéliser la file E2/M/1, c'est à dire une file unique comportant k clients (y compris celui en service) ; il y a un seul serveur (comme au 2°) ou en E.4) dont la durée de service suit la loi exponentielle de paramètre v ; le délai entre deux arrivées suit la loi Erlang-2 de paramètre u . Ce délai suit la même loi que le délai qui s'écoule entre deux arrivées de clients dans la situation considérée au 2°) ci-dessus. Autrement dit, l'introduction de l'automate considéré au 2°) a permis de **modéliser la loi Erlang-2 de façon markovienne**. On dit aussi qu'on a donné une modélisation markovienne de la loi Erlang-2 à l'aide **d'états fictifs** (les états 1 et 2 de l'automate). Le paragraphe F qui suit généralise ce point.

E.6 Automate exponentiel

1°) Définissons d'abord le modèle mathématique. L'ensemble E des états a deux éléments 1 et 2. Soit $u > 0$. On pose :

$$\begin{aligned} a(1, 2) &:= a(2, 1) := u \\ a(1, 1) &:= a(2, 2) := 0 \end{aligned}$$

2°) Ceci modélise un automate qui prend alternativement les états 1 et 2, le délai entre deux changements d'états suivant la loi exponentielle de paramètre u .

3°) Soit X_t l'état de l'automate à l'instant t .

Posons $x(t) := \text{Proba}[X_t = 1]$

$$y(t) := \text{Proba } [X_t = 2]$$

Les équations de Chapman-Kolmogorov s'écrivent :

$$x' = -u x + u y \quad \text{et} \quad y' = u x - u y$$

On en déduit facilement :

$$\text{Proba } [X_t = i \mid X_0 = j] = \frac{1}{2} [1 - e^{-2ut} + 2 \delta_i^j e^{-2ut}]$$

où δ_i^j est le symbole de Kronecker classique.

F. Etats fictifs

F.1 Processus non markovien

Reprenons l'exemple de la file E2/M/1 considérée en E.5-3°. Si on prend comme ensemble E' d'états, l'ensemble des entiers $k \geq 0$, k correspondant au nombre de clients dans la file, on a un processus non markovien : en effet, si X suit la loi Erlang 2, on a :

$$\text{Proba } [X \geq s+t \mid X \geq s] \neq \text{Proba } [X \geq t \mid X \geq 0] \quad .$$

Il faut alors se rappeler que le fait d'être markovien dépend de ce qu'on a choisi comme ensemble d'états : par exemple, **si, chaque état tient compte de tout le passé, le processus est markovien** ! Evidemment, mettre tout le passé dans l'ensemble des états est, en général, techniquement sans intérêt car ceci conduit à avoir un ensemble d'états beaucoup trop vaste. Par contre, on a souvent une solution intermédiaire qui consiste à "grossir" l'ensemble des états, par exemple en introduisant des états fictifs (cf. aussi, 3.E.3).

Par exemple, pour la file E2/M/1, si on prend comme ensemble d'états tous les couples (j,k) avec $j = 1$ ou $j = 2$ et $k \geq 0$ - comme cela est indiqué en E.5-1°) - c'est à dire que l'on introduit les états fictifs $j = 1$ et $j = 2$, alors on obtient un **processus markovien**.

F.2 Loi PH

Soit I un ensemble fini : en théorie on pourrait prendre I dénombrable mais, en pratique, puisqu'il faut pouvoir implémenter sur ordinateur, on suppose toujours que I est fini.

On se donne trois fonctions positives u, v et w, la fonction u étant définie sur $(I \times I)$, les fonctions v et w étant définies sur I et on suppose

$\sum_{j \in I} w(j) = 1$. Soit Z le processus markovien qui admet $(I \times \mathbb{N})$ comme ensemble d'états et tel que, pour tout couple (i, j) d'états, on a :

$$\begin{aligned} \lim_{h \downarrow 0} \frac{1}{h} \text{Proba} [Z_{t+h} = (j, k) \mid Z_t = (i, k)] &= u(i, j) \\ \lim_{h \downarrow 0} \frac{1}{h} \text{Proba} [Z_{t+h} = (j, k+1) \mid Z_t = (i, k)] &= v(i) w(j) \\ \text{Proba} [Z_0 = (j, 0)] &= w(j) \end{aligned}$$

La loi L du délai durant lequel ce processus Z reste dans l'ensemble $E_k := \{(i, k) : i \in I\}$ ne dépend pas de k .

Considérons un processus X qui admet E comme ensemble d'états. Pour alléger la présentation, on suppose que X est associé à un système S de files d'attente. On suppose que l'évolution de X est markovienne sauf que le délai entre deux arrivées de clients dans le système S suit la loi L . On peut alors rendre le processus markovien en prenant $(E \times I)$ comme ensemble d'états. La fonction u correspond aux changements d'états fictifs sans provoquer d'arrivée. Le paramètre $v(i)$ est le taux d'arrivée d'un client si on est dans l'état fictif i et $w(j)$ est la probabilité d'être dans l'état fictif j "juste après l'arrivée d'un client".

Evidemment, s'il y a k délais qui suivent des lois non exponentielles (au lieu d'un seul comme ci-dessus), on introduit k familles d'états fictifs pour rendre le processus markovien.

Quand une loi L est associée à une famille finie d'états fictifs comme indiqué plus haut, on dit souvent que c'est une loi de type PH, encore que les notations ne sont pas stabilisées dans ce domaine.

F.3 Loi hyperexponentielle

C'est le cas particulier de F.2 où la fonction u est identiquement nulle. La transformée de Laplace de la loi de L est alors :

$$\mathcal{L}(s) = \sum_{i \in I} w(i) v(i) / (s + v(i))$$

Supposer que l'on a une loi hyperexponentielle est une restriction assez forte concernant cette loi.

F.4 Modélisation avec pivot

1°) Un autre cas particulier consiste à supposer qu'il existe un élément i_0 de I tel que $w(i_0) = 1$ (et donc $w(j) = 0$ pour $j \neq i_0$). Autrement dit,

"immédiatement après la fin du délai", l'état fictif est dans l'état pivot i_0 . L'intérêt "technique" de cette propriété a été mise en évidence dans [Pel5]. En simplifiant, cette propriété permet d'éviter d'utiliser la technique de la "chaîne incluse" (sur laquelle nous reviendrons plus tard) ou, plus exactement, de la simplifier considérablement.

Il est également techniquement commode de supposer que l'on a $u(i,j) = 0$ pour $j \leq i$ (la matrice associée à u est triangulaire) : cette restriction est très faible.

2°) Par exemple, supposons que la loi à modéliser soit une combinaison convexe de loi de Cox "positive" (cf. F.5 ci-après), c'est à dire que la transformée de Laplace de cette loi peut s'écrire :

$$\mathcal{L}(s) = a(0) \frac{d}{s+d} + \sum_{i=1}^n a(i) \sum_{j=0}^{n(i)} c(i,j) \left\{ \prod_{k=0}^j \frac{b(i,k)}{s+b(i,k)} \prod_{k=0}^{j-1} [1-c(i,k)] \right\}$$

avec $\sum_{i=0}^n a(i) = 1$ et, quel que soit i , $b(i,0) = d$, $c(i,0) = 0$, $c(i,n(i)) = 1$ et $0 \leq c(i,j) \leq 1$.

On peut alors définir une modélisation avec pivot associée de la façon suivante. On prend comme ensemble I d'états fictifs l'ensemble constitué de l'élément 0 (l'état pivot) et des couples (i,j) avec $1 \leq i \leq n$ et $1 \leq j \leq n(i)$. On pose alors :

pour tout état fictif $e \neq 0$, $w(e) := 0$ et $w(0) := 1$;

$v(0) := a(0) d$, $v(i,j) := c(i,j) b(i,j)$

$u(0,(i,1)) := a(i) b(i,0)$

$u((i,k), (i,k+1)) := [1-c(i,k)] b(i,k)$

$u(e,e') := 0$ pour tous les autres couples (e,e') d'états fictifs.

F.5 Modélisation de Cox "positive" (cf. [Cox])

C'est un cas particulier de F.4. On pose $I := \{1, \dots, i, \dots, n\}$ et on prend 1 comme état pivot. On suppose que $u(i,j) = 0$ pour $j \neq i+1$. On pose :

pour $1 \leq i < n$, $x(i) := u(i,i+1)$ et $x(n) := 0$;

pour $1 \leq i \leq n$, $y_i := x(i) + v(i)$ et $z'(i) := x(i) / y_i$;

pour $1 \leq i \leq n$, $z_i := [1 - z'(i)] \prod_{j < i} z'(j)$.

La transformée de Laplace du délai associé à cette modélisation est :

$$\mathcal{L}(s) = \sum_{i=1}^n z_i \prod_{j \leq i} [y_j / (s + y_j)]$$

Le délai passé dans l'état fictif i suit la loi exponentielle de paramètre y_i ; une fois que ce délai est écoulé, le processus passe à l'état fictif $(i+1)$ avec la probabilité $z'(i)$ ou à l'état fictif 1 avec la probabilité $1-z'(i)$; un tel retour à l'état fictif 1 correspond à la fin du délai global.

Du point de vue théorique, l'écriture proposée en F.4 n'est pas plus générale que celle proposée en F.5 ci-dessus. Par contre, elle est plus souple et donc d'une utilisation technique plus simple.

F.6 Loi exponentielle "déguisée"

Quand on a à résoudre un problème où interviennent des délais dont la loi n'est pas exponentielle on commence, en général et dans un premier temps, par simplifier ce problème en supposant que tous les délais ont une loi exponentielle : l'espace des états est beaucoup plus petit (il n'y a pas d'états fictifs) et on dispose parfois de résultats analytiques.

Dans un deuxième temps on introduit une modélisation par états fictifs comme indiqué précédemment. Pour effectuer des vérifications par comparaison avec les résultats obtenus dans le premier temps (expliqué ci-dessus), il est alors souvent utile de noter que de telles modélisations sont équivalentes, en ce qui concerne les probabilités marginales, à la loi exponentielle pour certaines valeurs des paramètres.

Plus précisément, si, en F.2 (et donc dans les paragraphes suivants), on suppose que $v(i) = c$ est indépendant de i , **le délai considéré suit la loi exponentielle de paramètre c .**

G. Approximation de type CP

G.1 Introduction

Au paragraphe F qui précède on a donné les résultats essentiels du point de vue technique : notamment la formulation proposée en F.2 est à la fois simple, souple et générale.

L'objet du présent paragraphe G est de préciser les relations entre les diverses modélisations proposées au paragraphe F qui précède et, surtout, de montrer dans quelle mesure une loi de probabilité peut être approchée par une loi de type PH et par une loi de Cox "positive". **Les points essentiels de ce paragraphe sont donc les théorèmes G.4 et G.8 qui suivent.** Assez

curieusement et autant que l'auteur ait pu en juger, il n'y a pas d'étude "théorique" analogue publiée antérieurement autre que [Cox].

Or, dans [Cox], les problèmes d'approximation ne sont abordés que sous forme de questions. De plus, une lecture attentive montre **indiscutablement** que les transformées de Laplace des lois de service considérées dans [Cox] sont de la forme (cf. F.5) :

$$\mathcal{L}(s) = \sum_{i=1}^n z_i \prod_{j \leq i} [y_j / (s+y_j)]$$

la famille $(z_i)_{1 \leq i \leq n}$ étant une famille de réels **non nécessairement positifs**, ce qui est insuffisant pour une modélisation par états fictifs comme expliqué au paragraphe F qui précède. Plus précisément, il n'y a dans [Cox] aucun résultat analogue aux théorèmes G.4 et G.8 donnés ci-dessous (pour d'autres points de vue, cf. [Neu]).

G.2 Hypothèses et notations

Dans tout ce paragraphe G on considère un ensemble V de variables aléatoires positives dont chacune admet une densité de probabilité continue. De plus, on suppose que, pour tout triplet $(m, \sigma, \varepsilon')$ de réels strictement positifs, il existe un élément X de V de moyenne m' et d'écart-type σ' avec $|m-m'| \leq \varepsilon'$ et $\sigma' \leq \sigma$. Enfin, on suppose que toute combinaison convexe, finie ou infinie dénombrable, de telles densités est continue.

On notera V' (resp. W) l'ensemble des fonctions g (resp. intégrables au sens de Lebesgue) qui sont de la forme :

$$g = \sum_{j \in J} \lambda_j g_j$$

où $(\lambda_j, g_j)_{j \in J}$ est une famille, indexée par l'ensemble J fini (resp. dénombrable), de couples (λ_j, g_j) avec, quel que soit j , λ_j est un réel strictement positif et g_j est la densité de probabilité d'un élément de V .

G.3 Lemme préliminaire

Soit f une fonction définie, continue et positive sur \mathbb{R}^+ .

1°) Quel que soit $\varepsilon > 0$, il existe $\lambda > 0$ et un élément X de V tels que, si g est la densité de probabilité de X , on a :

$$\int_0^{+\infty} (\lambda g - f)^+(x) dx \leq \varepsilon \int_0^{+\infty} \lambda g(x) dx$$

où $y^+ = 0$ si $y \leq 0$ et $y^+ = y$ si $y \geq 0$.

2°) Quel que soit $\varepsilon > 0$, il existe un élément g de V' tel que :

$$\int_0^{+\infty} |g-f|(x) dx \leq \varepsilon$$

Preuve :

On se donne donc f et ε .

Soit $m > 0$ tel que $f(m) > 0$. Par continuité, il existe $\alpha > 0$ tel que $|x-m| < 2\alpha$ implique $f(x) > \beta$ avec $\beta := f(m)/2$. On pose $\varepsilon' := \alpha$, $\gamma := 1/\sqrt{\varepsilon}$ et $\sigma := \alpha\sqrt{\varepsilon}$.

Soit X un élément de V de moyenne m' et d'écart-type σ' avec $|m'-m| \leq \varepsilon'$ et $\sigma' < \sigma$. L'inégalité de Bienaymé-Tchebycheff s'écrit :

$$\text{Proba } [|X-m'| \geq \gamma \sigma'] \leq (1/\gamma)^2$$

Si g est la densité de probabilité de X et puisque $\sigma' \leq \sigma \leq (\alpha/\gamma)$, on a donc :

$$\int_A g(x) dx \leq \varepsilon \int_0^{+\infty} g(x) dx$$

où $A := \{x : |x-m'| \geq \alpha\}$.

Enfin, soit $\mu := \sup_{x \in A} g(x)$ et $\lambda := \beta/\mu$. Pour x n'appartenant pas à A , on a $\lambda g(x) \leq f(x)$, donc

$$\int (\lambda g-f)^+(x) dx \leq \varepsilon \int_0^{+\infty} \lambda g(x) dx$$

2°) On se donne toujours f et ε avec $0 < \varepsilon < 1$. On pose $\varepsilon' := \varepsilon/3$. Soit W' l'ensemble des éléments g de W tels que $\int_0^{+\infty} (g-f)^+(x) dx \leq \varepsilon' \int_0^{+\infty} g(x) dx$.

Dans W' on introduit la relation d'ordre $g R g'$ si $(g'-g)$ est un élément de W . En utilisant l'axiome du choix (cf. [Bou]), on sait qu'il existe une chaîne $(g_i)_{i \in I}$ d'éléments de W' maximale pour la relation d'ordre R . On pose :

$$s := \sup_{i \in I} \int_0^{+\infty} g_i(x) dx$$

Soit $(i(n))_{n>0}$ une suite croissante extraite de I telle que $s = \sup_{n>0} \int_0^{+\infty} g_{i(n)}(x) dx$.

On pose :

$$g'' := \sup_{n>0} g_{i(n)}$$

g'' est un élément de W (compte tenu de la relation d'ordre) et donc aussi un élément de W' (théorème de Lebesgue).

3°) Raisonnons par l'absurde et supposons que l'on n'ait pas $g'' \geq f$. La fonction g'' est continue sur le domaine $g'' \leq f$.

La fonction $(f-g'')^+$ étant continue, le 1°) permet d'affirmer qu'il existe $\lambda > 0$ et un élément X de V de densité de probabilité h telle que :

$$\int_0^{+\infty} (\lambda h - (f-g'')^+)(x) dx \leq \varepsilon' \int_0^{+\infty} \lambda h(x) dx$$

Or, on a :

$$(\lambda h + g'' - f)^+ = (\lambda h - (f-g'')^+)^+ + (g'' - f)^+$$

ce qui implique

$$\int_0^{+\infty} (\lambda h + g'' - f)^+(x) dx \leq \varepsilon' \int_0^{+\infty} (\lambda h + g'')(x) dx$$

La fonction $(\lambda h + g'')$ serait alors un élément de W' tel que, pour tout élément i de I , on aurait $g_i \leq R(\lambda h + g'')$ et $g_i \neq (\lambda h + g'')$, ce qui est impossible (puisque $(g_i)_{i \in I}$ est une chaîne maximale). Cette contradiction montre que la fonction $(f-g'')^+$ est nulle et on a :

$$\int_0^{+\infty} (g'' - f)^+(x) dx \leq \varepsilon' \int_0^{+\infty} g''(x) dx$$

Enfin, par construction de W , il existe un élément f' de V' tel que :

$$f' \leq g'' \text{ et } \int_0^{+\infty} (g'' - f')(x) dx \leq \varepsilon' \int_0^{+\infty} f'(x) dx$$

ce qui implique :

$$\begin{aligned} \int_0^{+\infty} |f' - f| (x) dx &\leq [\varepsilon' + \varepsilon'(1 + \varepsilon')] \int_0^{+\infty} f'(x) dx \\ &\leq \varepsilon \int_0^{+\infty} f'(x) dx \quad \text{c.q.f.d.} \end{aligned}$$

G.4 Théorème d'approximation

On considère les hypothèses et notations données en G.2. De plus, soit h une fonction définie sur \mathbb{R}^+ , positive et intégrable au sens de Riemann sur ce domaine et telle que :

$$\int_0^{+\infty} h(x) dx = 1 .$$

Autrement dit, h peut être considérée comme la densité de probabilité d'une variable aléatoire positive. Alors, quel que soit $\varepsilon > 0$, il existe un élément g de V' tel que :

$$\int_0^{+\infty} |g - h|(x) dx \leq \varepsilon$$

Preuve :

On suppose que h et ε sont donnés. Il existe une fonction f définie continue et positive sur \mathbb{R}^+ telle que

$$\int_0^{+\infty} |f - h|(x) dx \leq \varepsilon / 2 \quad \text{et} \quad \int_0^{+\infty} f(x) dx = 1$$

(puisque h est intégrable au sens de Riemann). Compte tenu du 2°) de G.3, il existe un élément g de V' tel que

$$\int_0^{+\infty} |f-g|(x) dx \leq \varepsilon / 2$$

ce qui implique

$$\int_0^{+\infty} |h-g|(x) dx \leq \varepsilon$$

G.5 Remarques

1°) On s'est limité au cas des variables aléatoires positives car c'est le cas important pour la suite. Il est clair que l'on aurait pu aussi bien considérer des variables aléatoires réelles quelconques ou des variables aléatoires à valeurs dans un intervalle B de l'axe réel (avec les modifications évidentes associées).

2°) De façon analogue, on pourrait prouver que V' est dense dans l'ensemble des fonctions positives pour des topologies autres que celle considérée ci-dessus.

G.6 Loi de type P.H.

1°) Nous allons d'abord introduire deux notations pour faciliter les explications ultérieures. Comme au paragraphe F, les diverses lois considérées seront caractérisées par leurs transformées de Laplace : ceci est une commodité technique. On pourrait tout aussi bien utiliser la transformée de Fourier ou même s'exprimer directement en termes de produits de convolution ou, équivalamment, en termes de mélanges de population et de variables indépendantes.

2°) Soit A un ensemble fini de réels strictement positifs (distincts) et m une fonction entière positive définie sur A . On pose :

$$\mu(A,m)(s) := \prod_{a \in A} \left(\frac{a}{s+a} \right)^{m(a)}$$

$\mu(A,m)$ est donc la transformée de Laplace du produit de convolution des lois d'Erlang de paramètre a (pour a parcourant A) et d'ordre $m(a)$.

3°) Soit I un ensemble fini. Pour tout élément i de I , soit $x(i)$ un réel positif, $A(i)$ un ensemble fini de réels strictement positifs (distincts) et $m(i)$ une fonction entière positive définie sur $A(i)$. On suppose que l'on a :

$$\sum_{i \in I} x_i = 1$$

On pose alors :

$$v((x_i, A_i, m_i)_{i \in I}) := \sum_{i \in I} x_i \mu(A_i, m_i)$$

Cette fonction est la transformée de Laplace d'une combinaison convexe (un mélange de populations) de lois du type introduit au 2°).

4°) Les lois qui admettent une telle transformée de Laplace sont parfois appelées de type PH. Notons d'abord qu'il y a évidemment équivalence entre la formulation proposée en F.2 et celle proposée dans ce 3°. La formulation proposée en F.2 est plus condensée et sera donc utilisée dans les chapitres ultérieurs.

Par contre, la formulation proposée dans ce 3°) est mieux adaptée à l'étude "théorique" des lois de type PH.

5°) Tout d'abord, cette formulation montre immédiatement qu'un mélange (i.e. une combinaison convexe) de lois de type PH est encore une loi de type PH. De même, un produit de convolution de lois de type PH est encore une loi de type PH.

6°) De plus, l'ensemble des lois d'Erlang satisfait évidemment aux hypothèses de l'ensemble V donné en G.2. Le théorème G.4 montre donc que, pour toute densité de probabilité f et pour tout $\varepsilon > 0$, il existe une loi de type PH dont la densité de probabilité g est telle que :

$$\int_0^{+\infty} |f-g|(x) dx \leq \varepsilon$$

On pourrait aussi se restreindre à une sous-famille de l'ensemble des lois de type PH, par exemple les lois dont la transformée de Laplace n'ait que des pôles simples ou n'ait que des pôles d'une forme particulière (par exemple 2^k avec k entier relatif).

7°) Enfin, on va maintenant montrer qu'une loi de type PH admet toujours une modélisation de Cox "positive" équivalente : ceci repose sur le lemme élémentaire qui suit.

G.7 Lemme sur les lois exponentielles

Lemme :

$$\frac{b}{s+b} = \frac{a}{s+a} \left\{ \frac{b}{a} + \frac{a-b}{a} \frac{b}{s+b} \right\}$$

Commentaires

La vérification de ce lemme est immédiate. Dans le cas où a et b sont

des réels avec $a > b > 0$, ce lemme signifie que la loi exponentielle de paramètre b est équivalente au mélange de la loi exponentielle de paramètre a (avec le coefficient b/a) et du produit de convolution des lois exponentielles de paramètres respectifs a et b (avec le coefficient $(1-b/a)$).

G.8 Théorème d'équivalence

Toute loi de type PH admet une modélisation de type CP (c'est à dire de Cox "positive").

Preuve :

1°) On considère donc la famille $(x_i, A_i, m_i)_{i \in I}$ comme définie en G.6, 3°) et la loi PH de transformée de Laplace $\forall (x_i, A_i, m_i)_{i \in I}$.

On pose $A := \bigcup_{i \in I} A_i$ et, pour tout élément a de A , $m(a) := \text{Sup}_{i \in I} m_i(a)$ puis :

$$n := \sum_{a \in A} m(a) .$$

Autrement dit, n est le nombre total de pôles, chaque pôle étant compté autant de fois que son ordre de multiplicité maximum. On se propose de prouver le théorème en raisonnant par récurrence sur n . On suppose donc que la propriété est vraie à l'ordre $(n-1)$.

2°) Pour tout élément i de I on pose :

$$a'_i := \text{Sup. } \{a : a \in A_i\}$$

$$A'_i := A_i \setminus \{a'_i\} \quad \text{si } m_i(a'_i) = 1$$

$$A'_i := A_i \quad \text{si } m_i(a'_i) > 1$$

$$m'_i(a'_i) := m_i(a'_i) - 1$$

$$m'_i(a) := m_i(a) \quad \text{si } a \neq a'_i$$

On pose aussi :

$$a' := \text{Sup}_{i \in I} a'_i$$

3°) Soit I'' l'ensemble des indices i tels que $\mu(A_i, m_i)$ est de la forme $a/(s+a)$ (c'est à dire que A_i est réduit à un seul élément a et que $m_i(a) = 1$). On pose $I' := I \setminus I''$.

Enfin, pour tout élément i de I , on pose $y(i) := a'_i / a'$ puis

$$z := 1 - \sum_{i \in I^n} x_i y_i$$

On note que $a' = a'_i$ si et seulement si a' appartient à A_i : dans ce cas, $1 - y_i = 0$. De plus, on peut supposer $z \neq 0$ (sinon I' est vide et le théorème est trivial).

4°) Le lemme élémentaire G.7 s'écrit

$$\frac{a'(i)}{s + a'(i)} = \frac{a'}{s + a'} \left\{ y_i + (1 - y_i) \frac{a'(i)}{s + a'(i)} \right\}$$

ce qui implique :

$$v((x_i, A_i, m_i)_{i \in I}) = \frac{a'}{s + a'} \{ (1 - z) + z Q \} \quad \text{où}$$

$$Q := \sum_{i \in I'} x_i y_i \mu(A'_i, m'_i) / z + \sum_{i \in I} x_i (1 - y_i) \mu(A_i, m_i) / z$$

$$\text{Or, } \sum_{i \in I'} x_i y_i / z + \sum_{i \in I} x_i (1 - y_i) / z = 1$$

donc Q est la transformée de Laplace d'une loi de type PH. De plus, le nombre de pôles de Q , chaque pôle étant compté autant de fois que son ordre de multiplicité maximum vaut $(n-1)$ (puisque a' est compté "une fois de moins") : l'hypothèse de récurrence donne alors que Q admet une écriture de Cox ; il en est donc de même de v , c.q.f.d.

G.9 Remarques

Le théorème G.8 qui précède montre que les modélisations PH (cf. F.2), avec pivot (cf. F.4) ou de Cox (cf. F.5) sont équivalentes quant aux lois que l'on peut ainsi atteindre. **Cela ne signifie pas que ces diverses modélisations sont équivalentes sur le plan technique.**

Notamment, si on construit un réseau de Cox suivant la procédure proposée dans la preuve de G.8, on commence par les états fictifs pour lesquels la loi exponentielle associée a le taux le plus grand : ces états vont donc avoir une probabilité faible, ce qui n'est pas satisfaisant sur le plan technique. La procédure proposée en G.8 doit donc être réservée aux études théoriques. Notamment, au 2°) de la preuve de G.8, on pourrait poser :

$$a'_i := \text{Inf. } \{a : a \in A_i\}$$

$$\text{puis } a' := \text{Sup. } a'_i \quad .$$

Il faudrait légèrement modifier la preuve proposée mais ce serait plus satisfaisant quant à la démarche technique. Toutefois, cela ne suffirait pas pour obtenir le "meilleur" réseau de Cox.

Nous allons maintenant donner quelques contre-exemples qui peuvent aider à mieux comprendre le modèle de Cox.

G.10 Contre-exemples

1°) Considérons quatre réels a, b, α et β avec $b > a > 0, \alpha + \beta b > 0, \alpha > 0$ et $\beta < 0$. On a :

$$u = v = w \text{ avec}$$

$$u := \alpha \frac{a}{s+a} + \beta \frac{b}{s+b}$$

$$v := \left(\alpha + \beta \frac{b}{a}\right) \frac{a}{s+a} + \beta \frac{a-b}{a} \frac{a}{s+a} \frac{b}{s+b}$$

$$w := \left(\beta + \alpha \frac{a}{b}\right) \frac{b}{s+b} + \alpha \frac{b-a}{b} \frac{a}{s+a} \frac{b}{s+b}$$

Compte tenu des inégalités imposées sur les coefficients, les deux écritures v et w correspondent à deux réseaux de Cox distincts (au sens strict c'est à dire avec des coefficients positifs).

Par contre, l'unicité de la décomposition d'une fraction rationnelle (et donc l'unicité de l'écriture telle que u) montre que u ne peut pas être la transformée de Laplace d'une loi hyperexponentielle.

En d'autres termes, ce contre-exemple montre qu'on peut avoir une loi qui admet plusieurs modélisations de Cox distinctes et qui n'est pas une loi hyperexponentielle.

2°) Soit a, b, c et ε quatre réels strictement positifs avec ε "petit". On a $u=v$ avec :

$$u := \frac{c}{s+c} \left\{ \frac{1}{2} + \frac{1}{2} \left[\varepsilon \frac{a}{s+a} + (1-\varepsilon) \frac{b}{s+b} \right] \right\}$$

$$v := \frac{c}{s+c} \left\{ \frac{1}{2} + \frac{1}{2} (1-\varepsilon + \varepsilon a/b) \frac{b}{s+b} + \frac{1}{2} \varepsilon (1 - \frac{a}{b}) \frac{a}{s+a} \frac{b}{s+b} \right\}$$

La loi associée à cette transformée de Laplace peut être modélisée, de plusieurs façons par trois états fictifs 1, 2 et 3. Plus précisément, l'écriture u peut être associée à la modélisation suivantes : l'automate se met d'abord dans l'état fictif 1 et y reste avec un délai qui suit la loi exponentielle de paramètre c . A la fin de ce délai, il y a une chance sur

deux que le service soit fini ; sinon, avec la probabilité (relative) ε , il va dans l'état fictif 2 et avec la probabilité $(1-\varepsilon)$ il va dans l'état fictif 3. Le délai passé dans l'état fictif 2 (resp. 3) suit la loi exponentielle de paramètre a (resp. b). A la fin de ce délai le service est terminé. C'est une modélisation où 1 est un "état pivot" (cf. F.4).

L'écriture associée à v peut être modélisée par un réseau de Cox avec les états fictifs successifs 1, 2 puis 3, le délai passé dans l'état fictif 1 (resp. 2,3) suivant la loi exponentielle de paramètre c (resp. b, a). Le service se termine à la sortie de l'état fictif 1 (resp. 2) avec la probabilité "relative" $\frac{1}{2}$ (resp. $(1-\varepsilon + \varepsilon a/b)$).

Ces deux modélisations correspondent à un même délai total. Pour autant, elles ne sont pas équivalentes au niveau des calculs ultérieurs ; notamment si a/b est "très petit" (de l'ordre de ε^2 par exemple), la modélisation de Cox est peu satisfaisante car l'état fictif 3, dont le rôle est crucial, ne peut être atteint que par l'intermédiaire de l'état fictif 2 dont la probabilité est très faible (puisque b/a est "très grand"). **La modélisation "avec pivot" proposée ci-dessus est alors, en général, beaucoup plus satisfaisante.**

Evidemment, le même contre-exemple, en supprimant l'état fictif 1, montre que la **loi hyperexponentielle peut, dans certains cas, être plus satisfaisante que la loi avec pivot associée.**

Chapitre 2

Régime Stationnaire**A. Coefficients de contraction****A.1 Préliminaires**

Dans une première étape, on aborde l'étude des processus en régime stationnaire par quelques propriétés mathématiques des **produits de matrices positives**. La nécessité de cette étude apparaîtra par la suite à plusieurs niveaux. Le sens "physique" de l'expression "**en régime stationnaire**" sera expliqué au paragraphe C.

On suppose que le lecteur connaît les propriétés et conventions usuelles en théorie des matrices. Notamment, une matrice réelle $(m \times n)$ a est une famille $(a_{i,j})$ de réels avec $1 \leq i \leq m$ et $1 \leq j \leq n$; $a_{i,j}$ est le terme de la i -ième ligne et de la j -ième colonne.

Si $m = n$, on dit que la matrice est carrée.

Si a est une matrice $(m \times n)$ et b une matrice $(n \times p)$, $a b$ est la matrice $(m \times p)$ de terme général

$$a b_{i,j} := (a b)_{i,j} := \sum_{k=1}^n a_{i,k} b_{k,j}$$

Si x est une matrice uniligne (resp. unicolonne), on pose $x_i := x_{1,i}$ (resp. $x_i := x_{i,1}$).

Etant donné a matrice $(n \times n)$, x matrice $(1 \times n)$, avec $x \neq 0$, et c réel, on dit que x est un vecteur propre de a associé à la valeur propre c si $x a = c x$.

La matrice a est dite positive (resp. strictement positive), si, pour tout couple (i,j) , on a $a_{i,j} \geq 0$ (resp. $a_{i,j} > 0$). Rappelons aussi qu'une matrice a est dite stochastique (cf. 1.C.3) si a est une matrice positive carrée $(n \times n)$ telle que, quel que soit i , $1 \leq i \leq n$, $\sum_{j=1}^n a_{i,j} = 1$.

La transposée de la matrice a est la matrice d définie par, quels que soient i et j , $d_{i,j} := a_{j,i}$.

A.2 Ecart angulaire

Pour tout couple (x,y) de matrices positives $(1 \times n)$ on pose :

$$\delta(x,y) := \sup_{i,j} (|x_i y_j - x_j y_i| / (x_i y_j + x_j y_i))$$

avec la convention $0/0 = 0$.

On dira que $\delta(x,y)$ est l'"**écart angulaire**" entre x et y . On a $\delta(x,y) \leq 1$. De plus $\delta(x,y) = 0$ implique que x et y sont proportionnels.

A.3 Coefficient de contraction

Soit m une famille de matrices positives $(n \times n)$, stable pour le produit matriciel. On dira que μ est un **coefficient de contraction** (on dit aussi coefficient d'ergodicité) - relativement à la distance angulaire δ - si les propriétés suivantes sont satisfaites :

- (i) μ est une fonction définie sur M et à valeurs dans $[0,1]$
- (ii) a et b éléments de M impliquent $\mu(a,b) \leq \mu(a) \mu(b)$
- (iii) x et y matrices positives $(1 \times n)$ et a élément de M implique $\delta(xa, ya) \leq \delta(x,y) \mu(a)$
- (iv) a matrice strictement positive implique $\mu(a) < 1$

L'idée consistant à considérer des coefficients de contraction est ancienne. La définition la plus connue est celle de coefficient de contraction de Birkhoff (cf. [Bir]) : on peut, par exemple, en trouver une étude détaillée dans [Sen]. La définition que nous proposons au paragraphe A.4 qui suit ne coïncide pas avec celle de Birkhoff et certaines propriétés du coefficient de contraction de Birkhoff ne sont pas satisfaites par la fonction μ définie en A.4. Par contre, notre définition conduit, nous semble-t-il, à une étude techniquement un peu plus simple.

A.4 Définition de μ

Lemme préliminaire : Soit $(p_i)_{i \in I}$ et $(q_i)_{i \in I}$ deux familles de nombres positifs indexées par le même ensemble fini I . On a :

$$(\sum_{i \in I} a_i) / (\sum_{i \in I} b_i) \leq \sup_{i \in I} (a_i / b_i)$$

Ce lemme élémentaire se vérifie en raisonnant par récurrence croissante sur $\text{card}(I)$ et en indexant par $\{1, \dots, n\}$ en sorte que $(a_i / b_i) \leq (a_{i+1} / b_{i+1})$.

Proposition : Soit M l'ensemble des matrices positives $(n \times n)$. Pour tout élément a de M on pose :

$$\mu(a) := \sup_{i,j,u,v} \{ |a_{i,u} a_{j,v} - a_{i,v} a_{j,u}| / (a_{i,u} a_{j,v} + a_{i,v} a_{j,u}) \}$$

Alors μ est un coefficient de contraction (au sens donné en A.3).

Preuve :

1°) Soit a et b deux éléments de M ; on a :

$$\mu(a b) := \sup_{i,j,u,v} F \quad \text{avec}$$

$$N := a_{i,u} a_{j,v} - a_{i,v} a_{j,u}$$

$$D := a_{i,u} a_{j,v} + a_{i,v} a_{j,u} \quad \text{et}$$

$$F := |N| / D$$

On pose :

$$x_r := a_{i,r}, \quad y_r := a_{j,r}, \quad c_r := b_{r,u}, \quad d_r := b_{r,v}.$$

On a :

$$N = \sum_{r,s} (x_r c_r y_s d_s - x_r d_r y_s c_s)$$

$$2N = \sum_{r,s} (x_r c_r y_s d_s - x_r y_s c_s d_r)$$

$$+ \sum_{r,s} (x_s y_r c_s d_r - x_s y_r c_r d_s)$$

$$2N = \sum_{r,s} (x_r y_s - x_s y_r)(c_r d_s - c_s d_r)$$

On pose :

$$t := (r,s), \quad f_t := x_r y_s - x_s y_r, \quad g_t := c_r d_s - c_s d_r$$

$$f'_t := x_r y_s + x_s y_r, \quad g'_t := c_r d_s + c_s d_r$$

ce qui précède s'écrit alors :

$$2N = \sum_t f_t g_t. \quad \text{On a de même}$$

$$2D = \sum_t f'_t g'_t \quad \text{ce qui implique}$$

$$F \leq (\sum_t |f_t g_t|) / (\sum_t f'_t g'_t)$$

soit, compte tenu du lemme préliminaire,

$$F \leq \sup_t \{ (f_t g_t) / (f'_t g'_t) \}$$

$$F \leq \sup_t (f_t / f'_t) \cdot \sup_t (g_t / g'_t)$$

$$\mu(a, b) \leq \mu(a) \mu(b)$$

2°) Soit x et y deux matrices positives $(1 \times n)$ et b un élément de M . On suppose que $(x / \|x\|) = 1$ et $(y / \|y\|) = 1$. On a :

$$\delta(xb, yb) = \sup_{u, v} (|N| / D) \text{ avec}$$

$$N := x b_u y b_v - x b_v y b_u$$

$$D := x b_u y b_v + x b_v y b_u$$

N et D ont donc les mêmes valeurs qu'au 1°) qui précède, ce qui donne la même inégalité finale. Or dans ce 2°), $\sup_t (f_t / f'_t) \leq \delta(x, y)$ ce qui prouve :

$$\delta(xb, yb) \leq \delta(x, y) \mu(b)$$

3°) Tous les coefficients étant positifs, on a

$$|a_{i,u} a_{j,v} - a_{i,v} a_{j,u}| \leq a_{i,u} a_{j,v} + a_{i,v} a_{j,u}$$

ce qui prouve $\mu(a) \leq 1$. De plus, l'inégalité précédente ne devient une égalité que pour $a_{i,v} a_{j,u} = 0$ ou $a_{i,u} a_{j,v} = 0$: on a donc $\mu(a) < 1$ pour toute matrice a strictement positive.

A.5 Matrices stochastiques

Considérons le cas où M est l'ensemble des matrices stochastiques $(n \times n)$. Dans ce cas, il peut être intéressant de définir différemment l'écart angulaire δ' et le coefficient de contraction associé μ' . Plus précisément, si x est une matrice $(1 \times n)$, on pose : $\|x\| := \sum_{i=1}^n |x_i|$. Puis, si x et y sont deux matrices positives, on pose :

$$\delta'(x, y) := \|(x / \|x\| - y / \|y\|)\|$$

Si a est une matrice stochastique $(n \times n)$ on pose :

$$\mu'(a) := \sup_{i, j} \left\{ \sum_{u, v} \frac{1}{2} |a_{i,u} a_{j,v} - a_{i,v} a_{j,u}| \right\}$$

On vérifie alors, comme en A.4, que μ' est un coefficient de

contraction relativement à δ '.

B. Produit de matrices positives

B.1 Vecteur propre

1°) Soit a une matrice positive ($n \times n$) telle que $\mu(a) < 1$ où μ est un coefficient de contraction associé à δ . Soit y une matrice uniligne ($1 \times n$). Soit $y_k := y a^k$; puisque l'écart angulaire entre y_k et y_{k+1} est inférieur à $\mu(a)^k$, la "**suite de Cauchy**" y_k converge en direction et, à la limite, on a $x a = c x$ avec c réel positif.

Ce vecteur propre x est le seul vecteur propre positif puisque $x = x a$ et $y = y a$ implique $\delta(x, y) \leq \delta(x, y) \mu(a)$ donc $\delta(x, y) = 0$; x ne dépend donc pas de y . Chaque ligne de a^k **converge "en direction"** vers x .

De plus, $\delta(x, y_k) \leq \mu(a)^k$ (**convergence géométrique** et ceci quelle que soit l'initialisation y).

Enfin, si $x a = c x$ (avec x et c positifs, c réel) et si $\delta(y, y a) = \varepsilon$ (avec y positif), on a $\delta(x, y) \leq \varepsilon / [1 - \mu(a)]$ (étude classique relative à la série géométrique) : on sait donc avec quel niveau de précision y peut être considéré comme une approximation de x .

2°) Soit a une matrice ($n \times n$) strictement positive.. On a $\mu(a) < 1$ et on peut utiliser le 1°) qui précède. Soit x une matrice positive ($1 \times n$) (non nulle) et c un réel positif tels que $x a = c x$. La matrice x est strictement positive puisque, quel que soit i , $\sum_j x_j a_{j,i} > 0$. Soit y une matrice réelle telle que $y a = c y$. En prenant h assez petit, la matrice $z = x + h y$ est positive et $z a = c z$. L'unicité prouvée au 1°) implique que z , et donc y est proportionnelle à x . Autrement dit, le **sous-espace propre associé à c est de dimension un**.

B.2 Matrices variables

En B.1 on a considéré le cas où on faisait le produit (k fois) de la matrice a par elle-même. Nous aurons aussi besoin (cf. § 6.B) du cas où on considère un produit de matrices distinctes.

Soit donc $(a_k)_{k \geq 0}$ une suite de matrices positives ($n \times n$). On pose

$\mu_k := \mu(a_k)$. On suppose que :

$\prod_{k=1}^{\infty} \mu_k = 0$ (ce qui équivaut à $\sum_{k=1}^{\infty} (1 - \mu_k) = +\infty$ suivant un théorème classique

d'analyse qu'on vérifie facilement en utilisant la fonction logarithme).

Soit $(b_k)_{k \geq 0}$ la suite de matrices définie par récurrence par $b_1 = a_1$ et $b_{k+1} = b_k a_{k+1}$.

On a $\lim_{k \rightarrow \infty} \mu(b_k) = 0$ donc, quel que soit la matrice positive uniligne

$(1 \times n) y$, $y b_k$ converge "en direction" vers une matrice uniligne x . Tout ce qui a été dit en B.1 s'étend au cadre considéré ici avec des modifications évidentes. Notamment le vecteur x est unique "en direction".

B.3 Matrice irréductible

Définition : Soit a une matrice positive $(n \times n)$. On dit que a est **irréductible** si, pour chaque couple (i, j) d'indices, $1 \leq i \leq n$, $1 \leq j \leq n$, il existe un entier $m := m(i, j)$ tel que $(a^m)_{ij} > 0$.

Compte-tenu des règles de calcul des produits matriciels ceci équivaut à dire que, pour tout couple (i, j) d'indices, il existe une séquence $(h(k))_{0 \leq k \leq m}$ d'indices telle que $h(0) = i$, $h(m) = j$ et, quel que soit k , $0 \leq k < m$, $a_{h(k), h(k+1)} > 0$.

Nous verrons par la suite que cette notion de matrice irréductible est une **notion fondamentale** pour deux raisons complémentaires. D'une part, en pratique, on peut (presque) toujours se ramener au cas où l'on considère des matrices irréductibles. D'autre part, une matrice irréductible n'admet qu'un seul vecteur propre positif (à une constante multiplicative près). Plus précisément, on a :

Proposition : Soit a une matrice positive $(n \times n)$ irréductible. Il **existe une et une seule matrice positive** $(1 \times n) x$ satisfaisant aux deux conditions suivantes :

$$(i) \quad \sum_{i=1}^n x_i = 1$$

$$(ii) \quad \text{il existe un réel } c \text{ positif tel que } x a = c x.$$

De plus, le sous-espace propre associé à c est de **dimension un** et x est strictement positive. Enfin, si a est stochastique, $c = 1$.

Preuve :

$$1^\circ) \text{ Posons } b := e^a = \sum_{k=0}^{\infty} a^k / k! \quad (\text{cf. 1D4})$$

La matrice b est strictement positive ; il existe donc (cf. B.1) une matrice $(1 \times n)$ y positive unique telle que $y b = c' y$ (c' réel positif) et $\sum_{i=1}^n y_i = 1$. De plus, cette matrice y est strictement positive.

2°) Unicité : soit x satisfaisant aux propriétés (i) et (ii). On a $x b = e^c x$ donc $x = y$ et $c = \text{Log } c'$, c'est à dire que x et c sont uniques.

3°) Existence : inversement, posons $z := y a$; on a $z b = c' z$ (puisque $a b = b a$) donc z est proportionnel à y (puisque b n'admet qu'un seul vecteur propre positif à une constante multiplicative près). On a donc prouvé que y est un vecteur propre de a , c'est à dire que $x = y$ satisfait aux propriétés (i) et (ii).

4°) $x a = c x$ et $y a = c y$ impliquent $x e^a = e^c x$ et $y e^a = e^c y$, donc x et y sont proportionnels (cf. B.1).

5°) Supposons, de plus, que a soit stochastique (cf. 1.C.3) et soit x et c satisfaisant à (i) et (ii). On a :

$$c = \sum_{i=1}^n c x_i = \sum_{i=1}^n (x a)_i = \sum_{j=1}^n x_j \sum_{i=1}^n a_{j,i} = \sum_{j=1}^n x_j = 1$$

Remarques :

1°) Les paragraphes A et B qui précèdent ne prétendent évidemment pas résumer tous les résultats concernant les produits de matrices "finies" positives : pour plus de détails, dans le cas fini ou infini, on peut, par exemple, consulter [Rev], [Sen].

Par contre, nous verrons que ces paragraphes A et B résument l'essentiel de ce dont on a besoin pour la plupart des applications. Le cas des matrices positives non irréductibles sera évoqué au chapitre 7.

2°) Plus précisément, mis à part l'étude des produits de matrices positives qui sera utilisée aux chapitres 5, 6 et 7, le **résultat essentiel** pour la suite de ce cours est l'**unicité de la solution stationnaire quand la matrice considérée est irréductible** ; il est plus simple - et peut-être plus révélateur de la réalité expérimentale - de prouver cette unicité à l'aide des coefficients de contraction que par l'intermédiaire de la très belle théorie de Perron - Frobenius (cf., par exemple, [Gan] ou [Sen]) ; c'est pourquoi cette théorie n'est pas résumée dans ce cours.

3°) Soit a une matrice stochastique. Dans la démonstration qui précède, on a étudié le système $x a = x$ (a matrice, c réel positif) par l'intermédiaire du système $y e^a = e y$: cette astuce mathématique a une signification physique très précise. Si on observe un processus qui évolue à temps continu, le délai entre deux transitions suivant la loi exponentielle de

paramètre 1 et, pour chaque transition, la matrice associée à cette transition étant a , alors la loi stationnaire y de ce processus est telle $y e^a = e y$ et on a $x = y$ (cf. 3.E.1).

C. Modélisation et régime stationnaire

C.1 Introduction

A partir de ce paragraphe C, le chapitre 2 utilise fortement les notions introduites dans le chapitre 1. Plus précisément, on considère un processus markovien $(X_t)_{t \in T}$ dont l'ensemble des états E est fini ou dénombrable. On se propose d'étudier ce processus quand t tend vers l'infini. En fait, on considérera essentiellement le cas où E est fini.

Du point de vue concret, cela signifie que l'on étudie le processus X sur un intervalle de temps **assez long** pour qu'il y ait un "**grand nombre**" de **transitions**. Par ailleurs, on supposera X homogène, c'est à dire qu'on étudie X sur un intervalle de temps **assez court** pour que les paramètres qui régissent l'évolution de X **varient peu**.

C.2 Exemples et remarques

1°) Reprenons les divers exemples introduits aux paragraphes 1.A et 1.E. Dans le cas de la marche aléatoire (cf. 1.A.2), quels que soient $\varepsilon > 0$ et $n > 0$, il existe un entier $m(\varepsilon, n)$ tel que, quel que soit $k \geq m(\varepsilon, n)$ on a $\text{Proba}[|X_k| \leq n] \leq \varepsilon$. Autrement dit, pour toute partie finie A de l'ensemble des états, la probabilité pour le processus d'être dans A à l'instant k (au " k -ième kilomètre parcouru") tend vers zéro quand k tend vers l'infini. On dit parfois que la "masse (associée à la probabilité) part à l'infini". Dans ce cas l'étude stationnaire n'a pas d'intérêt. La même remarque vaut pour le processus de Poisson (cf. 1.A.3 et 1.E.3).

2°) Pour l'exemple considéré en 1.A.4 b) (cf. aussi 1.E.1), $\lim_{t \rightarrow \infty} \text{Proba}[X_t = 1] = 1$; on dit que le point 1 (état de panne) est un point absorbant. De même, dans le cas de la loi d'Erlang- n (cf. 1.E.2), $\lim_{t \rightarrow \infty} \text{Proba}[X_t = n] = 1$ (quelle que soit la condition initiale) : n est un point absorbant. Dans ces deux cas, l'étude stationnaire n'a pas d'intérêt.

Pour l'exemple considéré en 1.C.5, il y a des points absorbants ($Y_k = 0$ et $Y_k \geq b^{2^{m+1}}$) mais, comme on l'a déjà noté, on s'intéresse plutôt au comportement moyen du régime transitoire (cf. 3.D.2 et le chapitre 6).

3°) Considérons maintenant le cas de la file M/M/1 (cf. 1.E.4).

Pour $v \geq u$, $\lim_{t \rightarrow \infty} \text{Proba}[X_t = k] = 0$ (cf. le 1° qui précède). Par contre,

pour $u < v$, on peut prouver que $\lim_{t \rightarrow \infty} \text{Proba}[X_t = k] = q_k$, la famille $(q_k)_{k \geq 0}$

étant telle que $\sum_{k=0}^{\infty} q_k = 1$ et, quel que soit $k \geq 1$, $v q_k = u q_{k-1}$. Cette limite ne dépend pas des conditions initiales.

Pour $v < u$, on dit que la famille $(q_k)_{k \geq 0}$ est la famille des **probabilités stationnaires**.

4°) Si on étudie un guichet de 14H à 18H, ce guichet satisfaisant aux hypothèses de la file M/M/1, et si le nombre moyen d'arrivées est d'un client à l'heure, le régime stationnaire n'a pas d'intérêt. Par contre, supposons que l'on étudie, un "bus" de liaison à l'intérieur d'un central téléphonique ou d'un ordinateur, les divers paramètres étant supposés fixes (ou "à peu près fixes") de 14H à 18H ; supposons qu'en première approximation on puisse modéliser ce bus à l'aide d'une file M/M/1, la moyenne de la durée de service étant une nanoseconde : il est évident que, dans ce cas, **du point de vue concret, l'essentiel est l'étude en régime stationnaire**.

5°) Il y a évidemment des exemples d'études en régime stationnaire pour lesquels il n'y a pas lieu d'utiliser les techniques proposées dans ce cours (cf. [Pel-3]).

C.3 Cas d'une chaîne homogène

1°) On considère donc le cas où $T = \mathbb{N}$; on pose :

$$p_{i,j} := \text{Proba}[X_{k+1} = j \mid X_k = i]$$

et on suppose que cette quantité ne dépend pas de k (cas homogène). On a vu (cf. 1.C.) que la matrice p est stochastique et que :

$$\text{Proba}[X_k = j \mid X_0 = i] = (p^k)_{i,j}$$

2°) On suppose que p est irréductible ; soit q la matrice positive $(1 \times n)$ vecteur propre de p et telle que $\sum_{i=1}^n q_i = 1$ (cf. B.3).

On dit que q est la probabilité stationnaire associée à la matrice p .

Soit c la constante positive telle que $q p = c q$. On a vu en B.3 que $c = 1$.

On a donc $q p^k = q$ c'est à dire que :

$$\{ \forall i, \text{Proba}[X_0 = i] = q_i \} \text{ implique } \{ \forall i, \forall k, \text{Proba}[X_k = i] = q_i \}$$

3°) Supposons que p soit irréductible et que, pour tout couple (i,j) , la

$$a'(v,u) = a(v,u) / s [1+n(u)] \text{ et } q'(v) = w q(v) / s [1+n(u)]$$

Dans les deux cas (1°) et 2°), on a donc :
 $q'(v) a'(v,u) = w q(v) a(v,u)$.

Finalement, on arrive à :

$$\sum_{v \in W(u)} q'(v) a'(v,u) = w \sum_{v \in W(u)} q(v) a(v,u)$$

Cette relation et 4D7 reportées dans 4D5 donnent exactement la relation 4D6.

Interprétation

Le cas le plus simple est le cas où la relation 4D5 est la balance locale (pour une classe ou pour la réunion de plusieurs classes) relativement à une station A de R ; dans ce cas, les termes $\tau(\overset{\uparrow}{u}, \overset{\uparrow}{A})$ et $\tau(\overset{\downarrow}{u}, \overset{\downarrow}{A})$ ne font intervenir les coefficients $a(u,v)$ que pour des états v tels que $n(v) \geq n(u)$. Les hypothèses données en D.4 sont donc satisfaites.

Notamment, si, dans un réseau R, il y a balance locale pour les stations S et A, on peut modifier les taux de service dans S (comme indiqué précédemment) sans perdre ces propriétés de balance locale.

E. Produit de deux réseaux

E.1 Introduction

Ce paragraphe E est indispensable si on veut disposer d'un cadre suffisamment souple et général au niveau des applications.

L'idée essentielle de ce paragraphe est la suivante : on considère un réseau R' de files d'attente (en un sens très général) ; on suppose que ce réseau R' possède une station S' pour laquelle il y a balance locale par classe. Alors, on peut remplacer cette station S' par tout un réseau R'' sous réserve que le "noeud" de liaison entre R' et R'' puisse être considéré, relativement à R'', comme une station S'' pour laquelle il y a balance locale par classe (dans R'').

Quand on effectue cette substitution, la probabilité stationnaire du réseau composé (R',R'') est, à une constante multiplicative près, le produit des probabilités stationnaires associées à R' et R'' respectivement. De plus, s'il y a balance locale pour la station A, autre que S', dans R', il y a encore balance locale pour A dans le réseau composé (R',R'') : on peut donc, à nouveau, remplacer la station A par tout un réseau, etc... ; autrement dit,

C.4 Cas à temps continu

On considère un processus homogène $(X_t)_{t \in T}$ avec $T = \mathbb{R}^+$ et dont l'ensemble $E := \{1, \dots, n\}$ des états est fini. On pose :

$$a_{i,j} := \lim_{h \downarrow 0} \{\text{Proba}[X_{t+h} = j \mid X_t = i] / h\} \quad \text{et} \\ g_t(i, j) := \text{Proba}[X_t = j \mid X_0 = i]$$

On a vu en 1.D.5 que $g_t(i, j) = (e^{bt})_{i,j}$ où b est la matrice définie par

$$b_{i,j} := a_{i,j} \text{ si } i \neq j \text{ et } b_{i,i} := - \sum_{j=1}^n a_{i,j}.$$

On suppose que la matrice a est irréductible (cf. B.3). La matrice e^{bt} est une matrice stochastique strictement positive (cf. 1.D.4 et 8.A). Soit μ le coefficient de contraction défini en A.4. Pour $t > k$, avec k entier, on a :

$$\mu(e^{bt}) \leq (\mu(e^b))^k \quad \text{donc} \quad \lim_{t \rightarrow \infty} \mu(e^{bt}) = 0 \quad \text{et on a une convergence "géométrique".}$$

Comme en C.3.5°) qui précède (cf. B.1), chaque ligne de e^{bt} converge vers le seul vecteur propre positif q de e^b . Or on a :

$$g'_t(i, j) = -g_t(i, j) \sum_{k=1}^n a_{j,k} + \sum_{k=1}^n g_t(i, k) a_{k,j}$$

Puisque $g_t(i, j)$ converge vers q_j quand t tend vers l'infini, $g'_t(i, j)$ tend vers zéro et, à la limite, on a :

$$0 = \lim_{t \rightarrow \infty} \left\{ -g_t(i, j) \sum_{k=1}^n a_{j,k} + \sum_{k=1}^n g_t(i, k) a_{k,j} \right\}$$

ce qui donne, quel que soit $j \in E$:

$$q_j \sum_{k=1}^n a_{j,k} = \sum_{k=1}^n q_k a_{k,j}$$

Cette famille de relations - que l'on appellera les équations d'équilibre - est absolument fondamentale pour toute étude en régime stationnaire à temps continu. En fait, une très grande partie de la suite de ce cours va consister à étudier le système linéaire ainsi introduit.

D. Les équations d'équilibre

D.1 Hypothèses de base

Compte tenu de l'importance des équations que nous venons de mettre en évidence en C.4, nous allons en rappeler le support.

On considère un ensemble E : dans ce chapitre 2, on s'est limité au cas où E est infini dénombrable.

On se donne une fonction positive a définie sur $(E \times E)$. On appellera équations d'équilibre (associées à a) la famille suivante de relations dans laquelle les $(q_u)_{u \in E}$ sont les "inconnues" :

quel que soit u élément de E ,

$$q_u \sum_{v \in E} a_{u,v} = \sum_{v \in E} q_v a_{v,u}$$

Si $E = \{1, \dots, n\}$ et si on pose, pour tout couple (u, v) , $d_{u,v} := \delta_{u,v} \left(\sum_{w \in E} a_{u,w} \right)$

où $\delta_{u,v}$ est le symbole de Kronecker, les équations d'équilibre peuvent s'écrire matriciellement comme suit :

$$q d = q a$$

En fait, on s'intéresse surtout aux solutions q qui sont des (matrices) positives ; de plus, quel que soit le réel c positif, si q est une solution, $c q$ est aussi une solution. On va donc considérer les solutions positives q telles que $\sum_{u \in E} q_u = 1$. On dira alors que q est une **solution normalisée**.

D.2 Cas des processus

Résumons maintenant la relation entre **les équations d'équilibre** et la **loi limite d'un processus** (cf. aussi D.6).

On considère donc un processus markovien homogène $(X_t)_{t \in T}$ avec $T = \mathbb{R}^+$ et qui admet E comme ensemble d'états. On suppose que, pour tout couple (u, v) d'éléments de E , avec $u \neq v$, on a :

$$a_{u,v} := \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba}[X_{t+h} = v \mid X_t = u] \right\}$$

pour tout élément u de E on pose $a_{u,u} := 0$. Si la matrice a est irréductible, quelle que soit la loi de X_0 , pour tout élément u de E , la fonction $t \rightarrow \text{Proba}[X_t = u]$ admet une limite q_u quand t tend vers l'infini et la famille $(q_u)_{u \in E}$ (qui ne dépend pas de X_0) satisfait aux équations d'équilibre (telles que définies en D.1 ci-dessus).

En général, la façon la plus simple d'expliciter le système $q d = q a$ est d'écrire les équations de Chapman-Kolmogorov (cf. 1.D.2) puis d'écrire que les dérivées qui apparaissent dans ce système sont nulles, ce qui donne le système $q d = q a$. On conseille au lecteur novice d'adopter cette démarche pour retrouver le système $q d = q a$ dans les exemples qui sont étudiés par la suite. Une démarche heuristique plus brève sera proposée au chapitre 4.

D.3 Cas fini irréductible

Proposition : On suppose que E est un ensemble fini et on choisit $E := \{1, \dots, n\}$. Soit a une matrice positive $(n \times n)$. On suppose que a est **irréductible** (cf. B.3). Pour tout couple (u, v) d'éléments de E , on pose

$$d_{u,v} := \delta_{u,v} \sum_{w \in E} a_{u,w} .$$

Le sous-espace vectoriel des matrices q $(1 \times n)$ telles que $q d = q a$ est de dimension un ; il contient **un seul vecteur q normalisé**, c'est à dire tel que

$$\sum_{u \in E} q_u = 1 ; \text{ ce vecteur } q \text{ est strictement positif.}$$

Preuve :

1°) L'irréductibilité de a implique que, pour tout élément u de E , $d_{u,u} > 0$. La matrice d est donc inversible. Posons $x = q d$, ce qui équivaut à $q = x d^{-1}$, et $r = d^{-1} a$; le système $q d = q a$ équivaut au système $x = x r$.

2°) On vérifie immédiatement que la matrice r est stochastique et irréductible. La proposition D.3 est alors un corollaire de la proposition B.3.

D.4 Résolution de $q d = q a$

1°) Comme on l'a déjà noté, l'essentiel de la suite de ce cours porte sur la résolution du système $q d = q a$. En fait, on supposera (presque toujours) que la matrice a est irréductible.

Du point de vue "**théorique**", le **problème est alors résolu** compte tenu de la proposition D.3 : il y a une et une seule solution.

En fait, notamment quand on considère des processus markoviens, $n = \text{card}(E)$ est en général "très grand" : un tout petit réseau de files d'attente a facilement 10^{30} états. La résolution, même approchée, du système $q d = q a$ pose donc des problèmes "techniques" qui conduisent très vite à de vrais problèmes mathématiques.

2°) La résolution des systèmes linéaires fait l'objet d'une littérature abondante dont on considère qu'elle relève, en général, de l'"analyse

numérique". Evidemment, quand on veut résoudre le système $q d = q a$, on a très souvent à utiliser une technique générale de résolution : ce point sera repris au chapitre 5.

3°) Une technique classique en analyse numérique est l'utilisation de méthodes itératives, notamment pour calculer le vecteur propre associé à la valeur propre de module maximal d'une matrice.

On a noté (en D.3) que le système $q d = q a$ équivaut au système $x = x r$ avec $r = d^{-1} a$; soit r' une matrice positive qui commute avec r (i.e. $rr' = r'r$) et telle que $\mu(r') < 1$ (cf. A.3). Si y est vecteur positif tel que $c y = y r'$ avec c réel positif, y est proportionnel à x .

Résoudre le système $q d = q a$ équivaut donc à résoudre $c y = y r'$, y positif. On peut, par exemple, procéder comme indiqué en B.1 et étudier la limite en direction, c'est à dire en renormalisant à chaque itération, de $y(r')^k$.

On peut prendre $r' = \sum_{k=0}^j r^k$, j étant judicieusement choisi. Il faut toutefois noter qu'une telle matrice r' a beaucoup moins de termes nuls que la matrice initiale r . Ceci est avantageux pour certaines techniques itératives ($\mu(r')$ est plus petit) : par contre c'est un inconvénient si on veut utiliser des techniques liées aux "matrices creuses" (c'est à dire les matrices dont beaucoup de termes sont nuls)... La méthode itérative associée à B.1 n'est donc utilisée que pour de "petits" réseaux.

4°) On a noté en D.3 que le système $q d = q a$ équivaut au système $x = x r$; il équivaut aussi au système $q = q s$ avec $s = a d^{-1}$. Contrairement à ce qu'on pourrait croire, ces trois systèmes ne sont pas équivalents au niveau de la "technique de calcul". Par exemple, étudier $x = q d$ au lieu de q augmente le "poids" relatif d'états u qui ont un taux d_u d'être quitté élevé. Cela peut être un avantage si cela permet à x de rester dans les limites tolérées par l'ordinateur considéré : en effet, même pour de "petits" réseaux, on considère souvent des états dont la probabilité est inférieure à 10^{-100} et on ne peut pas éliminer de tels états, alors que, en général, les ordinateurs ne mettent pas en mémoire de telles quantités, à moins de travailler en double précision ou d'utiliser un calculateur arithmétique spécifique.

Par contre, étudier x peut être un inconvénient si cela donne trop de poids à des états dont la probabilité est faible et sans intérêt.

Ces quelques remarques préliminaires et élémentaires montrent qu'il n'y a pas "une bonne méthode" pour résoudre le système $q d = q a$: le choix de la méthode numérique dépend du problème considéré.

5°) Les méthodes classiques d'analyse numérique n'utilisent, en

général, qu'accessoirement, le fait que la matrice a est positive. On introduira, aux chapitres 5 et 6, des algorithmes pour lesquels cette propriété est cruciale.

6°) Au chapitre 7 nous introduirons un algorithme qui utilise à la fois le fait que la matrice a est positive et le fait que la somme des colonnes de la matrice $(a-d)$ est identiquement nulle.

7°) Dans certains cas très particuliers mais très importants en pratique, on sait déterminer la solution explicite du système $q d = q a$. Ce point sera abordé au chapitre 4.

D.5 Théorème ergodique en moyenne

Revenons au cas où on considère un processus markovien homogène $(X_t)_{t \in T}$. Dans les paragraphes précédents deux questions ont été posées et ont obtenu une réponse partielle :

1°) Est-ce que $\text{Proba}[X_t = u]$ admet une limite q_u quand t tend vers l'infini ?

2°) Est-ce que cette limite est indépendante de la loi de X_0 (la loi initiale) ?

Si chacune de ces deux questions admet une réponse affirmative, on dit qu'il y a **ergodicité** (en moyenne). Un **théorème ergodique** est un théorème qui donne des conditions suffisantes pour qu'il y ait ergodicité, en moyenne ou par trajectoires (cf. §E qui suit). La littérature sur le "théorème ergodique" est considérable et ce théorème admet un très grand nombre de variantes suivant les hypothèses considérées ainsi que le niveau de précision et le niveau de généralité auxquels on se situe. Dans les paragraphes qui précèdent C.4, D.2 et D.3 notamment, on a vu qu'il y a ergodicité si l'ensemble E est fini et si la matrice a associée à l'évolution de X est irréductible. En général, il y a deux (ou trois) types de conditions pour l'ergodicité.

- D'une part des conditions de "**non-explosion**" : on a déjà noté en C.2.3°) que, pour la file M/M/1 (cf. 1.E.4), il y a explosion (la probabilité ne "charge" que l'infini) si et seulement si $v \geq u$; évidemment, si l'ensemble E est fini, il n'y a pas explosion.

- D'autre part des conditions de "**mélange**" : à titre de contre-exemple, supposons que (F, G) soit une partition de l'ensemble E des états et que $a_{u,v} = 0$ pour tout couple (u, v) appartenant à $(F \times G)$ ou $(G \times F)$; il ne peut pas y avoir d'"échange" entre les états de F et ceux de G : il ne peut donc pas y avoir indépendance de la limite par rapport aux conditions initiales ; on dit parfois qu'une matrice a est "mélangeante" si son coefficient de contraction (cf. A.3) est strictement inférieur à 1.

- Enfin, s'il s'agit d'une chaîne de Markov à temps discret, il faut aussi qu'il n'y ait pas "périodicité" (cf. le contre-exemple donné en C.3.4°).

D.6 Processus stationnaire

1°) Comme en D.2, considérons le cas d'un processus $X := (X_t)_{t \in T}$ markovien homogène qui évolue à temps continu et qui admet E comme ensemble d'états et a comme matrice d'évolution ; on définit d comme en D.1 et soit q une matrice positive normalisée telle que $q d = q a$.

On vérifie immédiatement que, si $b := d - a$, $q e^{bt} = q$. On dit que le processus X est un régime stationnaire si, pour tout élément t de T , q est la loi de X_t .

2°) Par ailleurs, pour toute séquence croissante $(t(0)), \dots, (t(n))$ d'éléments de T , soit la loi de $(X_{t(0)}, \dots, X_{t(n)})$ associée au semi-groupe e^{bt} et telle que la loi de $X_{t(0)}$ soit q . On vérifie immédiatement que ceci définit un système compatible de probabilités conjointes (puisque, quel que soit t , $q e^{bt} = q$) : cf. 1.B.4. La loi associée est appelée loi stationnaire (ceci vaut pour $T = \mathbb{R}^+$ et pour $T = \mathbb{R}$).

Ceci peut être interprété "physiquement" de la façon suivante. Quand on considère le processus X sur un temps très long (cf. D.2), sa loi à l'instant t "converge" vers la loi limite q . Pour t assez grand, on peut supposer que cette loi limite est (presque) atteinte : le processus X est alors en régime stationnaire.

E. Théorème ergodique ponctuel

E.1 Introduction

Nous avons déjà rencontré en D.5 une notion de théorème ergodique. En fait, il existe une grande variété de "théorèmes ergodiques".

Le "théorème ergodique" qui est le plus souvent utilisé (explicitement ou implicitement) dans les applications (notamment en **simulation**) consiste à admettre que si on observe un processus sur un temps suffisamment long - c'est à dire qu'on observe "une trajectoire" - les "**moyennes**" observées (**relativement à cette trajectoire**) coïncident avec ce qu'on peut "**espérer en moyenne**" à chaque instant.

Toute formalisation rigoureuse de cette propriété se situe nécessairement à un haut niveau mathématique, ce qui dépasse très

nettement l'objectif de ce chapitre. Nous allons donc seulement, dans ce paragraphe E, préciser un peu quelques propriétés ergodiques couramment utilisées. Une formalisation simple de la notion "brute" de trajectoire sera donnée au paragraphe F et son utilité sera illustrée par la preuve de la formule de Little.

E.2 Exemple de base

Nous allons commencer par le cas le plus simple : nous verrons d'ailleurs, au paragraphe E.3 qui suit, que l'on peut souvent se ramener à ce cas.

On considère un processus markovien homogène $(X_t)_{t \in T}$ qui admet E comme ensemble des états et $T = \mathbb{R}^+$ comme ensemble des temps. On suppose que E est fini et que la matrice qui régit l'évolution de X est irréductible. Soit f une fonction réelle définie sur E.

Considérons d'abord le cas où le processus évolue en régime stationnaire, c'est à dire que la loi de X_0 (la loi initiale) est la loi stationnaire ; quel que soit $t \in T$, la loi de X_t est la loi stationnaire. $E[f(X_t)]$ est donc une quantité fixe. On a aussi :

$$E[f(X_t)] = \frac{1}{t} \int_s^{s+t} E[f(X_r)] dr = \lim_{t \rightarrow \infty} \frac{1}{t} \int_s^{s+t} E[f(X_r)] dr$$

On peut considérer cette même limite pour un processus dont la loi initiale est quelconque : on a vu précédemment que la loi de X_r a une limite quand r tend vers l'infini et cette limite - la loi stationnaire - ne dépend pas de la loi initiale.

Autrement dit, on peut poser

$$\mu_X(f) := \lim_{t \rightarrow \infty} \frac{1}{t} \int_s^{s+t} E[f(X_r)] dr$$

On conçoit bien que, en général, **on peut intervertir le symbole E et la limite considérée ci-dessus**, soit :

$$\mu_X(f) = E \left\{ \lim_{t \rightarrow \infty} \frac{1}{t} \int_s^{s+t} f(X_r) dr \right\}$$

évidemment, ceci suppose d'avoir défini $f(X_t)$ - ce que l'on n'a pas fait auparavant puisque le processus a uniquement été considéré "en loi". La formalisation mathématique associée sera évoquée au chapitre 10 : X est alors une application définie sur $(\Omega \times T)$ et $X(\omega, t)$ est la valeur de X à l'instant t pour la trajectoire ω .

Le "**théorème ergodique**" le plus utilisé (souvent implicitement) exprime que l'on a aussi :

$$\mu_X(f) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_s^{s+t} f(X_r) dr$$

Autrement dit, **en régime stationnaire**, la moyenne $E[f(X_t)]$ à un instant t est égale à la moyenne "**le long d'une trajectoire**" sous réserve d'étudier cette trajectoire sur un temps suffisamment long.

Si on se place dans le cadre mathématique évoqué plus haut (cf. chapitre 10) et si l'ensemble E des états est fini, on peut prouver que cette propriété est satisfaite pour "**presque toute trajectoire**" c'est à dire pour tout élément ω de $(\Omega \setminus R)$ où R est un ensemble de probabilité nulle : c'est ce qu'on appelle un **théorème ergodique ponctuel**.

E.3 Dédoublage de processus

On considère le même processus $(X_t)_{t \in T}$ qu'au paragraphe E.2 qui précède. Par contre, considérons, par exemple, la "fonction" f telle que $f(X_t)$ égale 1 si et seulement si $X_t = e$ et si le processus X était dans l'état e' juste avant d'être dans l'état e , ces états e et e' étant deux éléments fixés de E . Cette fonction f n'est pas de la forme indiquée en E.2, mais on peut se ramener à cette forme de la façon suivante :

On pose $E' := E \times \{0, 1\}$; soit $(X'_t)_{t \in T}$ le processus à valeurs dans E' dont la loi d'évolution est définie de la façon suivante :

$$a'(u', v') := \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba}[X'_{t+h} = v' \mid X'_t = u'] \right\}$$

et pour $i = 0$ et $i = 1$:

$$a'[(u, i), (v, 0)] := a(u, v)(1 - \delta_{u, e'})$$

$$a'[(u, i), (v, 1)] := a(u, v) \delta_{u, e'}$$

c'est à dire que l'état $(u, 0)$ (resp. $(u, 1)$) de X' correspond au fait que X est

dans l'état u et que, juste avant la dernière transition, X n'était pas (resp. était) dans l'état e' .

Il suffit alors de poser $f'(u,i) := i \delta_{u,e}$ pour que $f'(X'_t)$ corresponde exactement à la "fonction" f introduite au début de ce paragraphe E.3.

Il est bien clair que des constructions analogues à celle proposée dans cet exemple peuvent être effectuées dans des situations beaucoup plus générales. En fait, pour la plupart des applications où on considère une moyenne temporelle, on peut se ramener, à l'aide d'une telle construction, à étudier $f(X_t)$ où f est une fonction définie sur E .

Au lieu d'utiliser cette notion de dédoublement de processus (ou ses généralisations) on peut considérer des fonctions markoviennes plus complexes que f , ce qui déplace les difficultés techniques.

E.4 Variables de comptage

Dans les paragraphes E.2 et E.3 qui précèdent, on a considéré la valeur moyenne de $f(X_t)$ où f est une fonction définie sur E . On peut aussi être amené à considérer **le nombre moyen de fois où un évènement s'est produit**, par exemple le nombre moyen de fois où on a la transition qui fait passer de e à e' , e et e' étant deux éléments fixés de E . Plus généralement, soit F une partie de $(E \times E)$; posons : $v_X(F,s,t) :=$ nombre de fois où l'on a (X_{r-}, X_{r+}) élément de F pour $s < r \leq t$.

Evidemment, cette expression n'a un sens que si X est défini "par trajectoires" (de même que $\int_s^{s+t} f(X_r) dr$).

Si le processus markovien homogène $(X_t)_{t \in \mathbb{R}_+}$ admet un ensemble d'état fini et une matrice d'évolution irréductible, on peut poser :

$$v_X(F) := \lim_{t \rightarrow +\infty} \left\{ \frac{1}{t} v_X(F,s,s+t) \right\}$$

c'est à dire que cette limite existe et ne dépend ni de s , ni de la loi initiale.

F. Etude trajectorielle

F.1 Introduction

Nous allons maintenant donner un formalisme très simple qui suffit pour modéliser la (ou une) trajectoire d'un processus X qui évolue en temps continu et dont l'ensemble E des états est fini. Ce formalisme sera notamment repris au paragraphe 3.E.

Cette notion de trajectoire est cruciale à plus d'un titre. D'une part, quand un processus "incarné" évolue au cours du temps, il n'évolue pas en valeur moyenne mais suivant une trajectoire. D'autre part, une méthode très répandue pour étudier un phénomène est de simuler une ou plusieurs trajectoires et d'en déduire des propriétés en moyenne : nous reviendrons sur ce point au chapitre 3 qui suit. Enfin, même si on se place sur le plan purement mathématique, il y a de nombreuses propriétés dont la démonstration s'effectue naturellement "par trajectoires" : nous donnerons l'exemple de la formule de Little au paragraphe F.4 qui suit.

F.2 Trajectoire

Dans tout ce paragraphe F, on considère un ensemble fini E et une suite $(e_k, s_k)_{k \geq 0}$ avec, quel que soit $k \geq 0$, e_k élément de E et $s_k > 0$. Quel que soit $k \geq 0$, on pose

$$t_k := \sum_{j < k} s_j \quad (\text{on a donc } t_0 := 0)$$

Soit f la fonction définie sur \mathbb{R}^+ par $f(t) = e_k$ pour $t_k \leq t < t_{k+1}$; $f(t)$ représente l'état du processus à l'instant t ; à l'instant t_k , $k > 0$, le processus passe de l'état e_{k-1} à l'état e_k puis il reste dans cet état e_k durant un laps de temps s_k , etc... On dira que f est la trajectoire du processus.

Pour alléger les notations, pour tout couple (s, t) de réels positifs et pour toute fonction réelle x définie et "continue par morceaux" sur \mathbb{R}^+ , on pose :

$$m(s, t, x) := \frac{1}{t} \int_s^{s+t} x(r) dr$$

De même, pour toute fonction définie sur \mathbb{N} et pour tout couple (s, t) de réels positifs, on pose :

$$n(s,t,x) := \frac{1}{t} \sum_{k \in A} x(k) \quad \text{avec } A := \{k : s \leq t_k < t\}$$

F.3 Fonction presque markovienne

Pour tout entier $i > 0$, soit \mathcal{F}_i l'ensemble des fonctions g à valeurs dans E définies sur l'ensemble $\{1, \dots, i\}$: autrement dit $(g(1), \dots, g(i))$ est une séquence de i éléments de E .

Soit x une fonction réelle définie sur \mathbb{N} ; on dira que x est presque markovienne (associée à la trajectoire f) si il existe $i > 0$ et g élément de \mathcal{F}_i tels que, quel que soit $k > i$,

$$x(k) = g[f(k-i+1), f(k-i+2), \dots, f(k)]$$

autrement dit la valeur en k ("à l'instant t_k ") de la fonction x est une fonction g (fixe) des valeurs de f aux instants t_{k-i+1}, \dots, t_k .

Soit y une famille réelle définie sur \mathbb{R}^+ ; on dira que y est presque markovienne (associée à la trajectoire f) si il existe une fonction x réelle définie sur \mathbb{N} markovienne (comme définie ci-dessus) telle que

$$y(t) = x(k) \quad \text{pour } t_k \leq t < t_{k+1}$$

Dans tout ce qui suit, on suppose que la trajectoire satisfait aux deux propriétés suivantes :

quelle que soit la fonction x réelle définie sur \mathbb{N} (resp. \mathbb{R}^+) presque markovienne, on peut poser :

$$\begin{aligned} n'(x) &:= \lim_{t \rightarrow \infty} n(s,t,x) \\ (\text{resp. } m'(x) &:= \lim_{t \rightarrow \infty} m(s,t,x)) \end{aligned}$$

c'est à dire que cette limite existe et ne dépend pas de s .

En fait, ces quantités $n'(x)$ et $m'(x)$ sont surtout intéressantes si elles sont "intrinsèques", c'est à dire indépendantes de la trajectoire considérée : cf. le paragraphe E qui précède.

Attention : si on procède par simulation (cf. chapitre 3) et si on doit étudier des fonctions telles que x , l'ensemble des états doit être suffisamment riche (c'est à dire contenir suffisamment d'information sur le passé du processus) pour que x soit strictement markovienne relativement à cet ensemble d'états. Les fonctions que nous appelons presque markoviennes sont les fonctions que l'on peut rendre strictement

markoviennes de façon simple et en conservant un ensemble d'états fini (cf. la technique du dédoublement de processus évoquée en E.3). Par ailleurs, pour procéder par simulation, il faut se donner une loi d'évolution ce qui n'est pas pris en compte par la notion "brute" de trajectoire définie en F.2 ci-dessus.

F.4 Formule de Little

1°) La "formule" de Little admet de nombreuses variantes. Considérons le cas où le processus étudié X est un "système A de files d'attente" - cette expression étant à prendre dans un sens très général - et soit "une" (ou "la") trajectoire associée à ce processus comme formalisé en F.2. Soit B un sous-ensemble de A. On suppose que des "clients" rentrent et sortent de ce sous-ensemble B et que, pour chaque état du processus X, le nombre de clients dans B est parfaitement déterminé. Pour alléger la présentation, on suppose qu'il ne peut pas y avoir simultanément de départs et d'arrivées de clients dans B ; par contre, les arrivées et les départs peuvent s'effectuer par paquets. Evidemment, il peut y avoir toutes sortes d'évolutions du processus X autres que celles de clients rentrant en B ou sortant de B.

2°) Soit y la fonction (markovienne) définie par $y(t) = j$ si et seulement si j est le nombre de clients dans B quand l'état est $f(t)$: autrement dit, y est le nombre de clients dans B à l'instant t pour la trajectoire considérée ; $m'(y)$ est alors le nombre moyen de clients dans B ; si L est la variable aléatoire associée au nombre de clients dans B (pour la trajectoire considérée), on a $E(L) = m'(y)$. On suppose $y(0) = 0$.

3°) Soit x la fonction (presque markovienne) définie sur \mathbb{N} par $x(k) = j$ si et seulement si la transition (e_{k-1}, e_k) correspond à l'entrée de j clients dans B ; $n'(x)$ est alors le nombre moyen de clients qui entrent dans B par unité de temps ; s'il n'y a pas explosion, c'est à dire si le nombre de clients dans B reste fini, $n'(x)$ est aussi le nombre moyen de clients qui sortent de B par unité de temps. Posons $a := n'(x)$: en général a est un paramètre "connu".

4°) Numérotions les clients qui entrent dans B à partir de l'instant $t = 0$ (ce numérotage étant choisi arbitrairement quand plusieurs clients entrent en même temps). Soit $t > 0$. Soit d_i le délai durant lequel le i -ième client reste dans B, l'observation étant arrêtée à l'instant t . Soit I le nombre de clients arrivés avant l'instant t . On a :

$$\sum_{i=1}^I d_i = \int_0^t y(s) ds$$

Il s'agit là d'une **propriété ensembliste élémentaire tout à fait indépendante de la théorie des processus**. Cette égalité peut aussi s'écrire :

$$\frac{I}{t} \cdot \left(\sum_{i=1}^I d_i \right) / I = \frac{1}{t} \int_0^t y(s) ds = m(o, t, y)$$

Faisons tendre t vers l'infini : (I/t) converge vers a ; $m(o, t, y)$ converge vers $m'(y) = E(L)$; $(\sum_{i=1}^I d_i) / I$ converge donc aussi vers une quantité qui est le délai moyen qu'un client donné passe dans B ; autrement dit, si D est la variable aléatoire associée au délai qu'un client passe dans B (pour la trajectoire considérée), on a :

$$a \quad E(D) = E(L)$$

5°) Il n'est pas sans intérêt, même pour le praticien, de noter la force et la faiblesse de la démonstration qui précède : contrairement à ce qui est parfois exprimé, cette démonstration ne repose pas sur les propriétés des processus régénératifs mais simplement sur des propriétés de "valeurs moyennes" ("lois des grands nombres") ; elle est donc très générale ; par contre, la preuve qui précède ne donne aucune information sur la "vitesse" avec laquelle, par exemple, $m(o, t, y)$ converge vers $m'(y)$ - vitesse qui peut dépendre de la trajectoire considérée - ni sur le caractère intrinsèque (indépendant de la trajectoire étudiée) de $m'(y)$.

Enfin, que le lecteur ne se laisse pas abuser par certaines démonstrations, apparemment savantes, qui ne font que "noyer le poisson" au milieu d'un formalisme inadapté et qui, en fait, ne prouvent rien de plus que ce qui précède.

6°) En plus des hypothèses données au 1°), on suppose que la sortie de B est assurée par un serveur U unique, qui sert les clients un par un. Soit z la fonction (markovienne) définie par $z(t) := 0$ (resp. 1) si B est vide (resp. non vide) à l'instant t . On a évidemment :

$$t = \int_0^t z(s) ds + \int_0^t [1 - z(s)] ds \quad \text{soit}$$

$$1 = m'(z) + m'(1-z)$$

De plus, soit u_i la durée durant laquelle le i -ième client est servi par le serveur U , l'observation étant arrêtée à l'instant t . On a :

$$\int_0^t z(s) ds = \sum_{i=1}^I u_i$$

ce qui peut aussi s'écrire :

$$m(o, t, z) = \frac{1}{t} \cdot \left(\sum_{i=1}^I u_i \right) / I$$

quand t tend vers l'infini, $m(o, t, z)$ tend vers $m'(z)$, $(1/t)$ tend vers a (cf. 4°)

et $\left(\sum_{i=1}^I u_i \right) / I$ tend vers la durée de service moyenne de chaque client, c'est à dire vers $E(S)$ si S est la variable aléatoire associée à la durée de service d'un client. En général, $E(S)$ est un paramètre "connu". On a finalement :

$$m'(1-z) = 1 - a E(S)$$

où $m'(1-z)$ est la "probabilité pour que le serveur U soit inoccupé". Cette "formule" suscite évidemment les mêmes commentaires qu'au 5°).

Chapitre 3

Simulation

A Modélisation

A.1 Introduction

Un double but est assigné à ce chapitre. D'une part, on y donne les notions de base de la simulation informatique. D'autre part, on y fait constamment référence aux notions mathématiques définies aux chapitres 1 et 2. **Pour le praticien, ce chapitre 3 peut donc être considéré comme une introduction à l'ensemble de ce cours.**

Le paragraphe A précise les **modèles considérés**. La notion de "**suite de nombres au hasard**" est expliquée au paragraphe B : cette notion est indispensable pour toute simulation informatique. Au paragraphe C, la **simulation informatique** est située par rapport aux **autres simulations expérimentales**. Les techniques de base de cette simulation informatique sont introduites, à partir d'exemples, aux paragraphes D (cas d'une chaîne de Markov) et E (cas d'une évolution à "temps continu"). Une esquisse de conclusion est donnée au paragraphe F.

Dans ce chapitre 3, on suppose presque toujours que l'on étudie le **régime stationnaire d'un processus markovien homogène** ; certaines des techniques proposées peuvent être utilisées dans des contextes plus généraux mais, le plus souvent, on se ramène à ce cadre : en effet, le principe de base de la simulation consiste à simuler une ou plusieurs trajectoires sur un temps très long, et donc en régime stationnaire. Toutefois, il faut bien comprendre ce que signifie cette expression : cf. notamment l'exemple D.2 où on étudie le comportement "moyen" d'un régime transitoire ; de même, aux paragraphes E.3 et E.4 on étudie la file $GI/GI/1$: le processus associé est rendu markovien par l'introduction d'états fictifs ou parce que l'état à chaque instant prend en compte le délai passé depuis la dernière arrivée et depuis le début du dernier service.

A.2 Cadre markovien

L'essentiel de ce cours porte sur l'étude de phénomènes markoviens en régime transitoire (chapitre 1) ou stationnaire (chapitre 2) dont l'ensemble E des états est fini ou dénombrable : la première étape de la modélisation, y compris quand on procède par simulation, est donc de choisir un ensemble E d'états suffisamment "riche" pour que, relativement à cet ensemble d'états l'évolution du processus et les fonctions à étudier soient markoviennes.

La notion de processus markovien a été **définie** en 1.B.3 ; rappelons notamment que, si $(X_t)_{t \in T}$ est un processus markovien qui admet E comme ensemble d'états, l'évolution du processus **après** l'instant t ne dépend **que** de l'état du processus à l'instant t et **non pas** de ce qui s'est passé **avant** l'instant t . Rappelons aussi que cette propriété dépend du choix de E et révèle le caractère crucial de la notion de loi exponentielle (cf. 1.A.4). On a vu en 1.F qu'en "grossissant" E on pouvait prendre en compte des lois non exponentielles, le processus X à étudier étant quand même markovien. De même, on a vu en 1.C.5, 2.E.3, etc... que, si on veut étudier des fonctions y liées au passé du processus markovien X , il peut être nécessaire de "grossir" l'ensemble E en sorte que y **ne dépende que de l'état e à l'instant t** .

Avant toute modélisation ou simulation en univers stochastique évolutif, il est indispensable d'avoir assimilé l'interprétation "concrète" de ce que l'on vient de rappeler.

A.3 Arrivées poissonniennes

Prenons un exemple simple et bien connu : supposons (cf. 1.E.3) que l'on veuille étudier les appels qui arrivent à un standard S en provenance de n abonnés (par exemple $n = 10.000$). En général, on suppose que les divers abonnés sont indépendants les uns des autres : qu'est-ce que cela veut dire "concrètement" ?

Cela signifie que l'on suppose que le fait que l'abonné a essaye "d'avoir la ligne" ne modifie en rien la probabilité pour l'abonné b , distinct de a , de faire de même. En général, cette hypothèse est très bien satisfaite si a et b disposent de lignes distinctes les reliant au standard S .

Supposons qu'on étudie le standard S de 20 H à 23 H et que, à 22 H une panne d'électricité se produise dans le secteur correspondant au standard S : le nombre moyen de nouveaux appels par minute d'observation va varier : notamment il va fortement augmenter à 21H30 et 22H30 (le tarif étant plus faible) et à 22H (compte tenu des appels à destination d'E.D.F.) : **ceci ne remet pas en cause l'hypothèse d'indépendance précitée** ; par contre le "taux" d'appels par abonné (c'est à dire la probabilité qu'un abonné demande la ligne sur un bref intervalle de temps) évolue au cours du temps.

On peut donc raisonnablement supposer que le flux des nouveaux appels est un **flux poissonnien** de paramètre λ variable, du moins en première approximation. Evidemment, cette approximation est insuffisante si le poids relatif des abonnés "saturés" est important : il est bien connu que, s'il fait beau, les cabines publiques des bords de mer sont saturées tous les soirs du 15 juillet au 15 août : il y a un nouvel appel immédiatement après chaque fin de communication.

A.4 Répétition d'appels et caractère markovien

Au paragraphe A.3 qui précède on a bien précisé que le flot des "nouveaux appels" est poissonnien. Si le standard est proche de la saturation, des abonnés n'obtiennent pas la ligne et ils vont donc rappeler incessamment. Si l'ensemble E des états ne prend en compte que les appels acceptés par le standard, l'évolution ne peut pas être markovienne. Là encore, pour rendre l'évolution markovienne il faut - et cela peut ne pas suffire car il faut aussi prendre en compte l'évolution interne du standard - grossir l'ensemble E des états, l'état du processus prenant en compte le **nombre d'abonnés en instance de répétition d'appels** : suivant la finesse de la modélisation, on distingue ou non les abonnés en fonction du nombre de tentatives antérieures avortées.

De même qu'en A.3, cet exemple montre, s'il en était besoin !, que choisir une modélisation markovienne valable **nécessite une très bonne connaissance du "terrain"** c'est à dire du phénomène à modéliser.

A.5 Régime stationnaire

On a déjà noté en 2.C.2, 4°), que l'intérêt de l'étude du régime stationnaire dépend des exemples. Dans le cas du standard S considéré ci-dessus, le nombre de nouveaux appels (par exemple une dizaine par minutes) est grand eu égard à la variation globalement assez lente du paramètre a (même si la valeur de a est discontinue aux instants de changement de tarifs). Dans ce cas, **c'est l'étude stationnaire qui est de loin la plus cruciale**, la "montée en charge" étant très rapide.

Plus généralement, avant d'effectuer une modélisation précise d'un phénomène, il est indispensable de savoir si on peut, on non, se contenter de l'étude en régime stationnaire : en effet, des difficultés techniques de l'étude en régime transitoire sont beaucoup plus grandes que celles de l'étude en régime stationnaire (toutes choses égales par ailleurs) : c'est d'ailleurs pour cela que ce cours porte essentiellement sur l'étude en régime stationnaire.

A.6 Cas E infini

Considérons toujours le cas du standard S correspondant à n abonnés (n "grand"). Soit X_t le nombre d'abonnés qui ont ou veulent avoir la ligne à un instant t donné. Supposons que l'on soit en sursaturation, c'est à dire que le paramètre a associé aux nouveaux appels soit strictement supérieur au débit maximum d du standard. En première approximation, X_t suit la même loi qu'une file $M/M/1$ de taux d'arrivée a et de taux de départ d : on a noté que, pour une telle file, le régime stationnaire n'a pas d'intérêt puisqu'il ne charge que l'infini (cf. 2.C.2,3°)).

Cela signifie-t-il qu'il faut effectuer une étude détaillée du comportement "à l'infini" de la file M/M/1 ? Non, bien sûr !. Il faut simplement se rappeler que le flot d'entrée a été approché par un processus de Poisson : ici, cette approximation n'est pas valable et il faut revenir au modèle initial qui prend en compte les n abonnés.

Cette remarque évidente et de bon sens a un caractère très général: plus précisément, assez souvent, l'ensemble E des états est "naturellement" fini, ou, tout du moins, il faut se limiter à des modélisations pour lesquelles E est fini si on veut procéder (partiellement) par simulation informatique ou algorithme de calcul : il peut être commode, pour faciliter l'étude mathématique, de plonger E dans un ensemble infini : par exemple pour remplacer une loi binomiale par une loi de Poisson. **Il ne faut pas, pour autant, oublier le point de départ pour lequel E est fini.**

Ceci explique que, dans ce cours, le cas où E est infini n'est presque jamais étudié "pour lui-même" alors que ce cas E infini fait l'objet de dizaines d'ouvrages et de milliers d'articles. Par contre, mais ceci est un autre problème, il arrive fréquemment que le nombre d'éléments de E soit très grand. Or, assez curieusement, les méthodes utilisées jusques ici pour l'étude théorique du cas où E est infini **ne sont** en général **pas adaptées** au cas où E est fini mais très grand (cf. 7.D.6)..

B. Nombres au hasard

B.1 Définition "mathématique"

Pour toute **simulation informatique**, il faut disposer d'une **suite de nombres au hasard** : nous allons donc d'abord préciser cette notion.

Soit $m > 0$. Soit M l'ensemble des entiers i , $0 \leq i < m$. Soit $(x_k)_{k \geq 0}$ une suite d'éléments de M . Que signifie l'expression "la suite $(x_k)_{k \geq 0}$ est une suite de nombres au "hasard" ? Cette expression n'est pas purement mathématique. Cela veut dire que cette suite peut être considérée comme une **"réalisation du hasard" associée à une suite $(X_k)_{k \geq 0}$ de variables aléatoires**, chaque "tirage" X_k étant indépendant des précédents et la loi de chaque tirage X_k étant la loi équirépartie sur M . Notons aussi que la "suite" $(x_k)_{k \geq 0}$ n'est pas "infinie" contrairement à la définition mathématique d'une suite. La suite $(x_k)_{k \geq 0}$ doit donc satisfaire aux propriétés statistiques des suites de variables indépendantes à valeurs

dans M . Notamment, pour j fixé, $j \geq 1$, la suite $(y_k)_{k \geq 0}$ de séquences définies par $y_k := (x_{jk}, \dots, x_{jk+i}, \dots, x_{jk+j-1})$ peut être considérée comme une "réalisation du hasard" associée à une suite $(Y_k)_{k \geq 0}$ de variables aléatoires indépendantes, chaque variable Y_k suivant la loi équirépartie sur M^j .

B.2 Construction

On peut construire une suite de nombres au hasard par des moyens mécaniques. Par exemple, si $m=6$, on lance "indéfiniment" un dé à six faces et on pose $x_k = u_k - 1$ où u_k est le n° de la face associée au k -ième lancer. On peut aussi utiliser des appareils un peu plus sophistiqués qui relèvent eux-mêmes les nombres x_k , ceux-ci étant obtenus par des moyens mécaniques ou électroniques : par exemple, on lit le temps suivant une période voisine de la milliseconde la lecture étant effectuée avec une précision de 10^{-6} millisecondes et on pose $x_k =$ le dernier chiffre significatif associé à cette lecture. Si le délai D entre deux observations (de moyenne 1 milliseconde) suit une loi "raisonnable" avec un écart-type suffisant (par exemple, supérieur à 10^{-4} milliseconde), la suite $(x_k)_{k \geq 0}$ peut être considérée comme une suite de nombres au hasard avec $m=10$.

En fait, de plus en plus, les suites de nombres au hasard sont "construites" par des "**générateurs informatiques**", c'est à dire des programmes (ou sous-programmes) informatiques.

Actuellement, la technique de loin la plus utilisée est celle des **congruences mixtes**. Le principe en est le suivant : considérons d'abord le cas où m est grand ; on construit la suite x_k par récurrence suivant la formule $x_{k+1} = a x_k + b \pmod{m}$ c'est à dire que, x_k étant connu, on calcule $y_k = a x_k + b$ puis on effectue la division euclidienne de y_k par m ce qui donne $y_k = m z_k + x_{k+1}$ avec $0 \leq x_{k+1} \leq (m-1)$: **se donner un générateur informatique c'est tout simplement se donner les nombres a, b, m et x_1** .

Il y a évidemment de nombreuses études sur le choix des nombres a, b, m et x_1 : notamment, des conditions ont été données pour que la période de la suite $(x_k)_{k \geq 0}$ soit la plus grande possible (la suite $(x_k)_{k \geq 0}$ a

nécessairement une période inférieure ou égale à $(m+1)$ puisque, après $(m+1)$ itérations, on est sûr de retomber sur une valeur déjà obtenue) ; il n'est pas évident que ceci soit un bon critère statistique comme le montre le contre-exemple suivant : soit $(x_k)_{k>0}$ la suite définie par $x'_k = x_k$ et $x'_{2k+1} = d$; cette suite $(x'_k)_{k>0}$ a une période qui, en général, est double de celle de $(x_k)_{k>0}$ et, pour autant, c'est une bien mauvaise suite de nombres au hasard. Intuitivement, il est "probable" qu'il y a intérêt à ce que la période soit d'un ordre de grandeur nettement inférieur à m (par exemple $m/10$).

Si on connaît un ordre de grandeur, soit p , de la période on peut, après p itérations redémarrer avec une valeur initiale x_{p+1} obtenue à l'aide d'un autre générateur (informatique ou non) ce qui provoque un décalage et suffit pour "casser" la période (cette procédure est souvent appelée randomize). Ces quelques remarques élémentaires montrent que le choix d'un "bon" générateur informatique laisse encore beaucoup de place à ... disons l'intuition.

Par contre, il y a quelques règles qu'il faut évidemment observer. Notamment, b doit être premier avec m et il est conseillé de choisir a et b du même ordre de grandeur, $b = b'b''$ avec $b' < b''$ et $a=1$ (modulo b') (pour ne pas avoir de période inférieure à b') ou des propriétés analogues.

Enfin, dans tout ce qui précède on a supposé m grand ; pour perdre le moins possible d'information, il y a intérêt à choisir $m \geq 2^{n-2}$ si n est le nombre de bits attribués à l'écriture des entiers positifs par l'ordinateur considéré. On verra, en B.4, que l'on peut construire d'autres suites $(y_k)_{k>0}$ à partir de la suite $(x_k)_{k>0}$: notamment, on peut poser $y_k := x_k$ (modulo m') pour avoir $0 \leq y_k \leq m'$.

Notons aussi que, contrairement aux générateurs mécaniques ou électroniques, **le générateur informatique permet de reproduire plusieurs fois exactement la même suite $(x_k)_{k>0}$** ce qui peut être utile pour effectuer des vérifications.

B.3 Sondages et codages

Ce paragraphe B.3 sort quelque peu du cadre général de ce cours : néanmoins il nous semble utile de rappeler deux utilisations des générateurs de nombres au hasard autres que la simulation qui sera abordée au paragraphe C qui suit.

Mis à part les jeux de hasard (dès, cartes, roulettes, loteries, etc...),

les suites de nombres au hasard ont d'abord été utilisées pour effectuer les **sondages**. Le principe en est élémentaire : supposons que l'on veuille choisir "**au hasard**" j éléments parmi m' éléments distincts. On numérote (explicitement ou implicitement) les m' éléments à étudier puis on choisit j nombres au hasard (en général, si deux de ces nombres sont égaux, on effectue un tirage supplémentaire : c'est ce qu'on appelle le sondage "avec remplacement") à valeurs dans l'ensemble $\{1, \dots, m'\}$ et on sonde les j éléments associés à ces j nombres. Il y a d'ailleurs des livres de "nombres au hasard".

Une autre application des générateurs informatiques de nombres au hasard est le **codage** dont l'importance va évidemment aller en croissant. Là encore, le principe en est élémentaire ; pour autant, ce type de codage a une efficacité qui rend obsolètes les études de décodages basés sur des propriétés statistiques.

On considère donc un message. Ce message, même s'il s'agit d'une émission de télévision, peut être écrit sous forme d'une suite $(z_k)_{k \geq 0}$ de nombres, chaque nombre étant compris entre 0 et $(m-1)$. En fait, de plus en plus, tous les messages seront écrits de cette façon ce qui facilite leur transmission (puisque tous les messages sont de même type et que ce type se prête particulièrement bien aux découpages, vérifications, etc...) : c'est la transmission numérique (par "paquets").

Soit $(x_k)_{k \geq 0}$ une suite de nombres au hasard à valeurs dans $\{0, \dots, m-1\}$, suite que l'on peut reproduire à volonté (cas des générateurs informatiques). On pose $y_k = x_k + z_k \pmod{m}$ et on transmet la suite $(y_k)_{k \geq 0}$.

L'initié qui connaît la suite $(x_k)_{k \geq 0}$ n'a évidemment aucun mal à retrouver la suite $(z_k)_{k \geq 0}$ connaissant la suite $(y_k)_{k \geq 0}$; par contre, le non-initié qui ne connaît pas la suite $(x_k)_{k \geq 0}$ aura beaucoup de mal à retrouver la suite $(z_k)_{k \geq 0}$: notamment, il n'y a aucune corrélation statistique entre les suites $(y_k)_{k \geq 0}$ et $(z_k)_{k \geq 0}$ (si $(x_k)_{k \geq 0}$ est une "bonne" suite de nombres au hasard). En fait, dans ce cas, **le seul décodage viable consiste à découvrir le générateur** de la suite $(x_k)_{k \geq 0}$: on laisse au lecteur "en manque" de roman policier le soin d'imaginer de tels générateurs à la fois souples et faciles à protéger.

B.4 Tests

Nous avons noté au paragraphe précédent que, mis à part quelques règles simples, il n'y a pas de critères décisifs pour bien choisir les nombres a , b , m et x_1 . Par contre, on peut effectuer de **nombreux tests pour "vérifier"** qu'un générateur est satisfaisant.

Comme on l'a déjà noté, "autrefois", les suites de nombres au hasard servaient essentiellement à effectuer les sondages : par exemple, les livres de nombres au hasard donnent, en même temps, les propriétés statistiques essentielles des suites de nombres proposées et ceci pour chaque page : moyennes, variances, test du χ^2 , etc...

De très nombreux tests ont été imaginés pour étudier les "qualités" d'un générateur aléatoire : ces tests ne portent pas uniquement sur la loi d'occurrence de chaque nombre mais sur d'autres occurrences plus subtiles : occurrence de la séquence (u,v) , de la séquence (u,v,w) (avec u , v et w donnés), etc... et pour chaque tel type d'occurrences on peut effectuer les études classiques (moyenne, variance, test du χ^2 , ...).

En fait, les constructeurs sont en général très discrets sur les "qualités" des générateurs aléatoires qu'ils utilisent explicitement ou implicitement (cas des "langages de simulation"). Par ailleurs, les qualités souhaitables pour un générateur dépendent de l'utilisation projetée : certains types de générateurs de périodes courtes mais robustes et bien "mélangeants" peuvent être excellents pour le codage et peu satisfaisants pour certaines simulations.

B.5 Générateur et simulation

Le problème général de la simulation sera abordé au paragraphe C qui suit. Le but de ce paragraphe B.5 est simplement de montrer que l'on peut simuler **n'importe quelle suite de variables aléatoires** à l'aide d'une suite $(x_k)_{k>0}$ de nombres au hasard. On suppose donc que cette suite est donnée avec, quel que soit k , $0 \leq x_k < m$. Posons $x'_k := x_k/m$.

1°) Cas d'une variable discrète : supposons que l'on veuille utiliser x_k pour simuler la valeur de la variable aléatoire U à valeurs dans l'ensemble fini V ; soit $(v_i)_{1 \leq i \leq s}$ la famille ordonnée des éléments de V et $q_i := \text{Proba}[U = v_i]$; posons $f_i = \sum_{j \leq i} q_j$; soit i tel que $f_i < x'_k \leq f_{i+1}$; la réalisation du hasard pour U associée à x_k consiste à choisir $U = v_{i+1}$. On vérifie

immédiatement que (avec la précision associée à m) $\text{proba}[U=v_i] = q_i$. Evidemment, si la précision associée à m est insuffisante (par exemple s'il existe i tel que $q_i < (1/m)$), il faut utiliser le couple (x_k, x_{k+1}) au lieu du seul nombre x_k , ce qui revient à opérer en double précision à ce niveau.

2°) Cas d'une variable continue : supposons que l'on veuille utiliser x_k pour simuler la valeur de la variable aléatoire réelle U qui admet $F(t)$ comme fonction de répartition, F étant continue. On prend alors pour U la valeur $F^{-1}(x'_k)$ et on a, avec la précision associée à m , $\text{Proba}[U \leq t] = F(t)$. S'il est important de simuler des valeurs de U (notamment des valeurs très petites ou très grandes) telles que les probabilités associées soient faibles (par exemple inférieures à $1/m$) il faut, là encore, utiliser le couple (x_k, x_{k+1}) .

3°) On peut, de façon analogue, simuler des variables plus générales que celles évoquées au 1°) et au 2°) ci-dessus.

C. Simulation expérimentale

C.1 Un simulateur est markovien

Dans le monde moderne, la simulation a un rôle déterminant et ce rôle ne peut qu'aller en croissant. Les simulateurs pour l'apprentissage des cosmonautes, des pilotes d'avions ou des simples futurs conducteurs de voitures sont bien connus. De plus en plus souvent, la construction d'un appareil A est précédée de celle d'un appareil simplifié dont l'évolution est accélérée pour simuler et donc prévoir les réactions de A . Les accidents de la route sont simulés, etc...

Ce cours ne se propose pas de faire une étude générale de ces divers types de simulation, mais simplement de préciser le formalisme mathématique associé quand l'ensemble des états est fini. Notons que cette étude "mathématique" met en évidence quelques propriétés valables pour n'importe quelle simulation.

Notamment, un simulateur ne peut valablement évoluer que s'il dispose à chaque instant de tous les éléments qui déterminent l'évolution ultérieure : par exemple, un simulateur de vol ne peut pas se contenter d'avoir en mémoire l'état de la cabine de pilotage : il lui faut aussi la position, la vitesse et l'accélération de l'avion, etc. Autrement dit, l'ensemble des états possibles du simulateur doit être suffisamment riche pour que l'évolution, à partir de l'instant t , ne dépende que des réactions du pilote et de l'état du simulateur à l'instant t . En langage

mathématique, cela signifie que l'ensemble E des états du simulateur doit être suffisamment riche pour que, par rapport à E , l'évolution soit **markovienne** : **un simulateur doit être markovien**. Attention : dans le cas du simulateur de vol, c'est l'ensemble "simulateur + pilote" qui est markovien.

On peut donc appliquer à l'étude d'un simulateur tout ce qui a été dit précédemment sur les processus markoviens : nous allons reprendre une partie de cette étude sous ce nouvel angle. Plus précisément, dans ce qui suit, nous allons étudier l'évolution du simulateur. Par contre, nous n'aborderons pas le problème de l'adéquation du simulateur en tant que modélisation de la réalité. Cet aspect du problème ne doit pour autant pas être négligé : notamment il n'est peut-être pas nécessaire de chercher la "paille" dans le simulateur si on ne voit pas la "poutre" qui sépare le simulateur de la réalité. Notons que, pour bien concevoir le simulateur, il faut avoir compris quel ensemble d'états rend le modèle réel markovien : par exemple, si on veut simuler les appels à un standard (cf. A.4), le simulateur doit prendre en compte les répétitions d'appels.

Les informaticiens sont devenus et resteront l'une des composantes fondamentales du monde moderne. Il faut tout de même noter que certains d'entre eux (c'est l'exception !) ont un peu tendance à réinventer le "fil à couper le beurre". Au lieu de parler de processus markovien, est-il plus simple et plus convaincant de parler de principe de causalité, de principe de déterminisme, de condition de réalisabilité ou de condition de prédictabilité ?

C.2 E fini

1°) On se limite au cas où l'ensemble E des états est fini (éventuellement dénombrable) : en ce qui concerne les applications, c'est une restriction contraignante. Toutefois, une partie non négligeable des idées ou méthodes introduites dans ce cadre s'étend au cas d'un ensemble quelconque d'états. De plus, on peut se ramener, plus souvent qu'on ne le croit, au cas où E est fini à condition de considérer des ensembles E "très grands" : par exemple, l'ensemble des états possibles d'un ordinateur, si puissant soit-il, est **fini**, ce qui n'est pas sans conséquences "théoriques".

On a déjà noté en C.1 que l'ensemble E des états que l'on considère est l'ensemble des états du simulateur. Si le simulateur est un ordinateur, E est donc nécessairement fini.

2°) E étant fini, le formalisme mathématique associé - tout du moins quant à sa première étape de construction - a été donné en 1.B.3 ; on a signalé que, sauf pour quelques exemples élémentaires, il est impossible de donner une solution mathématique explicite des équations de Chapman-Kolmogorov (cf. 1.D.2) : dans le cas homogène, la solution donnée en 1.D.5 est une solution théorique. Par contre, il est possible de calculer la solution stationnaire (cf. chapitre 2) pour quelques modèles

simples : c'est l'objet de l'essentiel de ce cours.

C.3 Modélisation d'un simulateur

On considère un simulateur dont l'ensemble E des états est fini et qui évolue, à temps continu, de façon (partiellement) aléatoire mais markovienne. Par exemple, s'il s'agit d'un simulateur de vol, E est l'ensemble des états possibles du couple "simulateur + pilote", le caractère aléatoire étant essentiellement lié aux réactions du pilote. Pour un tel exemple, supposer que E est fini est très contraignant. De plus, si on ne s'intéresse qu'à l'appareil dont les réactions aux commandes sont déterminées et qui sert à simuler le vol sans vouloir prendre en compte les réactions du pilote, le formalisme présenté dans ce cours est inutile. Enfin, comme on l'a rappelé au paragraphe précédent, on a de solution utilisable que pour quelques modèles simples en régime homogène stationnaire. L'utilisation du présent cours pour modéliser un simulateur de vol se limite donc à des études sur les propriétés "stationnaires" de certaines réactions du pilote : états intermédiaires entre la veille et le sommeil, perception des divers messages, etc...

S'il s'agit d'une maquette ou d'un prototype, l'ensemble E des états est, en général, plus facile à cerner. Le caractère "aléatoire" est alors lié aux expériences effectuées : par exemple, s'il s'agit de tester un prototype d'autocommutateur, on va simuler l'environnement (cf. [Heb]) de façon partiellement aléatoire.

Pour un tel exemple, la mise au point du simulateur d'environnement suppose d'avoir bien compris ce qu'est un système markovien. De plus, l'essentiel est l'étude en régime homogène stationnaire : encore faut-il bien comprendre le sens de cette expression. Ce type de simulateur est donc susceptible de modélisation mathématique comme indiqué précédemment. Celle-ci peut apporter des compléments d'information et des "vérifications" mais, évidemment, dans l'état actuel des connaissances, elle ne saurait remplacer l'expérimentation du prototype même si le simulateur d'environnement est imparfait. En fait, **le principal intérêt de l'étude mathématique est de pouvoir être effectué avant la mise au point d'une maquette, laquelle "fige" l'essentiel de la structure de l'appareil.**

Actuellement, dans neuf cas sur dix (au moins !), cette étude "préliminaire" - qui, malheureusement, est souvent entreprise trop tard - est effectuée par "simulation informatique".

D. Simulation informatique d'une chaîne de Markov

D.1 Préliminaires

La simulation informatique est **le stade le plus avancé de la modélisation** théorique avant l'étude mathématique. Evidemment l'évolution doit être markovienne relativement à l'ensemble des états choisis : de plus, la loi de cette évolution doit être parfaitement déterminée pour pouvoir être implémentée sur ordinateur.

La simulation informatique présente deux avantages déterminants : d'une part, elle permet de prendre en compte **un très grand nombre d'états** ; d'autre part, elle est, en général, **relativement simple à programmer** : le plus difficile, et de loin, est de bien choisir un ensemble d'états et la loi d'évolution markovienne associée.

Nous allons d'abord considérer la simulation des chaînes de Markov à partir de deux exemples, le premier étant tout à fait élémentaire.

D.2 Un jeu élémentaire

Pour cet exemple, la solution stationnaire est relativement simple à calculer : l'intérêt de cet exemple est donc essentiellement d'expliquer la technique de base de la simulation informatique. Reprenons le cadre proposé en 1.C.5. C'est un cadre particulier où on s'intéresse essentiellement à l'étude des "points absorbants" $Y_k = 0$ et $Y_k \geq 2 \cdot 2^m$ et tout ce qui s'y rattache (probabilités associées, nombre de parties à effectuer avant d'atteindre ces points, etc...).

Soit f la fonction qui à x associe le logarithme en base 2 de (x/b) . On pose $X'_k = f(X_k)$ et $Y'_k = 2Y_k/b$. On met en mémoire les nombres X'_k et Y'_k et soit $(u_k)_{k>0}$ une suite de nombres au hasard, la variable aléatoire associée à chaque terme de cette suite suivant la loi équidistribuée sur l'ensemble $\{0, 1, \dots, 3r-1\}$: le nombre u_k donne l'issue de la k -ième partie, la partie étant perdue si $0 \leq u_k < 2r$ et gagnée si $2r \leq u_k \leq 3r-1$.

Autrement dit, si $0 < X'_k < n$, on pose $Y'_{k+1} = Y'_k$ et $X'_{k+1} = X'_k - 1$ si $0 \leq u_k < 2r$ et $X'_{k+1} = X'_k + 1$ si $2r \leq u_k < 3r$.

Si $X'_k = 0$ on pose $X'_{k+1} = 0$ et $Y'_{k+1} = Y'_k - 1$ si $0 \leq u_k < 2r$; au contraire, si $2r \leq u_k < 3r$, on pose $X'_{k+1} = 1$ et $Y'_{k+1} = Y'_k$.

Si $X'_k = n$ on pose $X'_{k+1} = n-1$ et $Y'_{k+1} = Y'_k$ si $0 \leq u_k < 2r$; au

contraire, si $2r \leq u_k < 3r$, on pose $X'_{k+1} = n$ et $Y'_{k+1} = Y'_k + 2^{n+1}$.

On arrête le programme quand $Y'_k = 0$ ou quand $Y'_k \geq 2^{m+2}$.

L'écriture informatique de cet "algorithme" est immédiate.

En fait, on s'intéresse au **comportement "moyen"** de ce qui peut se passer - ou va donc simuler plusieurs telles trajectoires **simultanément** (cas de processeurs travaillant en parallèle) ou **successivement** (dès que le programme précédent est terminé, on le "relance") avec toujours la même initialisation - puisque, pour cet exemple, elle est imposée - soit $X'_1 = 0$ et $Y'_1 = 2^{n+1}$.

On met également en mémoire tous les paramètres qui nous intéressent. Par exemple supposons que l'on s'intéresse, à la probabilité d'être ruiné : on garde en mémoire deux compteurs, l'un compte le nombre u de trajectoires qui se terminent par $Y'_k = 0$ et l'autre le nombre v de trajectoires qui se terminent par $Y'_k \geq 2^{m+2}$. Le rapport $u/(u+v)$ donne la probabilité (approchée) de finir ruiné. On peut aussi étudier le nombre de parties qu'on peut espérer jouer avant d'être ruiné, etc... On note que, dans cet exemple, on étudie en moyenne, c'est à dire en régime stationnaire, un comportement transitoire (cf. le chapitre 6) ! Notons aussi que la simulation informatique est tellement simple qu'on est amené à l'utiliser alors que la solution théorique exacte est accessible.

D.3 Bus avec blocages

Un bus assure la liaison entre ℓ émetteurs $(E_i)_{1 \leq i \leq \ell}$ et m récepteurs $(R_j)_{1 \leq j \leq m}$. A chaque rotation, le bus prend le message de tête dans un émetteur (s'il n'y a pas de message, il passe à l'émetteur suivant) et va le porter au récepteur auquel il est destiné. Si le récepteur n'est pas saturé, il prend le message. Sinon, le message est rendu à l'émetteur d'où il vient. A chaque rotation, le bus change d'émetteur et passe à l'émetteur suivant. Soit r_j la capacité du récepteur R_j .

On pose $L := \{1, \dots, \ell\}$ et $M := \{1, \dots, m\}$. Soit U l'ensemble des fonctions u définies sur L et à valeurs dans \mathbb{N} : $u(i)$ correspond au nombre de messages en attente dans l'émetteur i . Soit V l'ensemble des fonctions v définies sur L et à valeurs dans M : $v(i)$ correspond au récepteur auquel est destiné le message en tête de l'émetteur i . Soit W l'ensemble des fonctions w définies sur M telles que, quel que soit j , $1 \leq j \leq m$, $w(j) \leq r_j$: $w(j)$ représente le nombre de messages dans le récepteur R_j .

On pose $E := U \times V \times W \times L$; un élément e de E est donc un quadruplet

(u,v,w,i) : les interprétations de u , v et w ont été données précédemment et i correspond au numéro de l'émetteur qui sera servi au prochain passage du bus.

D.4 Les transitions

Une fois que l'ensemble des états est déterminé, il faut définir la loi des transitions. Ici, on se donne :

1°) une fonction réelle a définie sur $(L \times M)$; $a(i,j)$ est la probabilité pour un message dans l'émetteur i d'être destiné au récepteur j . On

suppose que, quel que soit i , $\sum_{j=1}^m a(i,j) = 1$.

2°) Une fonction réelle b définie sur $(L \times \mathbb{N})$; $b(i,k)$ est la probabilité qu'entre deux rotations du bus, il y ait k nouveaux messages dans l'émetteur i . On suppose que, quel que soit i , $\sum_{k \geq 0} b(i,k) = 1$.

3°) Une fonction réelle c définie sur $(M \times \mathbb{N})$; $c(j,k)$ est la probabilité qu'entre deux rotations du bus, il y ait k messages qui puissent être retirés du récepteur j : évidemment, si ce récepteur contient moins de k messages il est simplement vidé.

D.5 Le déroulement du programme

Le déroulement du programme est alors très simple. On part d'un état quelconque (par exemple, les émetteurs et les récepteurs sont vides) et on considère un très grand nombre de transitions avec, à chaque fois, les lois de probabilités adéquates.

Soit e l'état juste après la $(k-1)$ -ième transition (i.e. la $(k-1)$ -ième rotation du bus). Cet état étant "en mémoire" dans l'ordinateur, pour que le programme se déroule, il suffit de déterminer l'état e' après la k -ième transition. On détermine alors cet état e' en simulant, à l'aide d'une "suite de nombre au hasard", les quelques éventualités à choisir (comme cela est indiqué au paragraphe B qui précède).

Il n'est pas nécessaire d'être un spécialiste de la programmation pour voir que ce programme est facile à écrire.

Pour le déroulement expliqué ci-dessus, la taille mémoire nécessaire est très faible : elle est égale à $m \ell^2$ puisqu'il suffit de mettre en mémoire la valeur de e . Il faut aussi mettre en mémoire les informations que l'on veut garder.

Si on souhaite connaître la "probabilité stationnaire" pour tous les

états e de E , il faut, pour chacun de ces états e , mettre en mémoire le nombre d'occurrences de cet état au cours de la simulation : la taille mémoire nécessaire est alors égale à $\text{card}(E) = \ell \cdot m^\ell \cdot n^\ell \cdot r^m$ (si $r := \sup_j r_j$ et n le nombre de messages maximum par émetteur).

Mais, en général, au niveau des applications, on n'a pas besoin de la probabilité stationnaire complète mais seulement de certaines marginales : il suffit donc de mettre en mémoire le nombre d'occurrences associées à ces marginales.

Notons aussi que, si le compilateur utilisé est adapté au calcul "vectoriel" (i.e. peut travailler en parallèle), au lieu de simuler une trajectoire, on peut simuler n trajectoires, la loi d'évolution étant la même pour chaque trajectoire, les initialisations pouvant être, ou non, différentes. Evidemment, les générateurs informatiques doivent être distincts pour chaque trajectoire.

D.6 Interprétation mathématique

La mise en oeuvre du programme nécessite de connaître l'ensemble E des états et, pour tout couple (i,j) d'éléments de E , la probabilité $p_{i,j}$ de passer de l'état i à l'état j lors d'une rotation du bus. Soit $(q_i)_{i \in E}$ la solution stationnaire (cf. 2.C.3) associée. Sauf pour des valeurs aberrantes des paramètres, cette solution stationnaire positive est unique.

L'hypothèse de base de la simulation informatique (dans le cas discret) est d'admettre que, sous réserves d'effectuer une simulation suffisamment longue, la proportion d'occurrences de l'état i est voisine de q_i .

Autrement dit, la simulation informatique n'est rien d'autre qu'une méthode mathématique particulière pour déterminer la solution (unique) normalisée du système linéaire $q \cdot p = q$: comme on l'a déjà noté, au lieu de déterminer la solution complète q on se contente souvent de déterminer **certaines combinaisons linéaires** (les marginales utiles) de cette solution, c'est à dire certaines matrices $q \cdot r$ (où r est une matrice uni-colonne connue donnée).

La remarque évidente qui précède en appelle deux autres non moins évidentes. D'une part q est la solution d'un système linéaire avec tout ce que cela implique. D'autre part, si on cherche la solution normalisée du système $q \cdot p = q$, p étant une matrice stochastique, on peut toujours procéder par simulation informatique comme indiqué précédemment, même si le problème concret étudié ne correspond en rien à l'étude d'une évolution markovienne. Si on procède ainsi, on dit parfois qu'on utilise une **méthode de Monte Carlo**.

D.7 Viabilité de la méthode

Dans ce qui suit, on suppose que la modélisation proposée ci-dessus en D.3 et D.4 est bien adaptée au problème concret que l'on a à résoudre : pour autant, il ne faut pas oublier qu'il est souvent inutile d'effectuer une étude fine d'un modèle qui n'est qu'une grossière approximation de la réalité.

Pour ce modèle, où on a à résoudre $q \cdot p = q$, les méthodes mathématiques exactes sont (immédiatement) inutilisables sauf pour de toutes petites valeurs des paramètres ℓ , m , n et r : certes la matrice p est **une matrice creuse** (beaucoup de ses termes sont nuls : cf. 5.C) mais elle a une structure complexe et $\text{card}(E)$ est égal à 10^{31} si les quatre paramètres ℓ , m , n et r sont égaux à 10 (et il s'agit d'un modèle simple). De plus, ce système n'a pas de solution explicite simple.

Par contre, sauf pour des valeurs aberrantes des paramètres, il n'y a pas d'état "peu accessible" dont la probabilité est faible et le rôle crucial (contrairement, par exemple, aux études usuelles de fiabilité). Plus généralement, la matrice p^k est bien mélangeante dès que k est assez grand (par exemple $k = 10\ell$) : son coefficient de contraction (cf. 2.A) est nettement inférieur à 1. C'est donc **un exemple type** pour lequel la **simulation informatique donne d'excellents résultats**.

Cela n'interdit pas d'effectuer des études mathématiques portant sur des modèles simplifiés dont on peut tester la validité... par simulation.

E. Processus évoluant à temps continu

E.1 Cas d'un processus markovien homogène

Lorsque le phénomène étudié évolue à temps continu, il y a plusieurs techniques de simulation informatique qui se ramènent évidemment toutes à une étude à temps discret : à chaque pas, l'ensemble des possibilités informatiques dont on dispose est grand mais fini !

Nous allons commencer par le cas où la modélisation est la plus avancée du point de vue théorique. Plus précisément, dans ce paragraphe E.1, on suppose qu'on a su choisir un ensemble E d'états fini relativement auquel la loi d'évolution est **homogène markovienne et "connue"**. Comme en 1.D, pour tout couple (u,v) d'éléments de E , $u \neq v$, on pose :

$$a(u,v) := \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba} [X_{t+h} = v \mid X_t = u] \right\}. \text{ On pose aussi } a_{u,u} := 0.$$

Comme en 2.D.3 on pose $d_{u,v} := \delta_{u,v} \sum_{w \in E} a_{u,w}$.

La **probabilité stationnaire** est la "matrice" uniligne q telle que $q d = q a$ et $\sum_{u \in E} q_u = 1$. On suppose que a est irréductible (cf. 2.B.3) : q est alors unique (cf. 2.D.3).

Si on pose $r = d^{-1} a$, on a déjà noté en 2.D.3 que r est une matrice stochastique et que $q d = q a$ équivaut à $x = x r$ si $x = q d$. On peut alors résoudre le système $x = x r$ par simulation comme indiqué au paragraphe D qui précède.

Une interprétation concrète de ceci consiste à dire que, si on est dans l'état u , on accélère ou on ralentit le temps de façon à ce que la moyenne du délai du temps passé dans cet état u avant la prochaine transition soit égale à l'unité de temps, et ceci quel que soit l'état u (cette modification de l'échelle des temps dépend donc de l'état du processus).

Après avoir effectué cette modification de l'échelle des temps, on obtient un processus dont le générateur infinitésimal est associé à r : autrement dit, le délai entre deux transitions quelconques suit la loi exponentielle de paramètre 1. La probabilité stationnaire à temps continu est alors la même que la probabilité stationnaire de la chaîne de Markov associée à l'observation du processus immédiatement après chaque transition (cf. la remarque 3°) dans 2.B.3).

E.2 Discrétisation du temps

La technique proposée au paragraphe E.1 qui précède est, en général, la plus "**économique**" en ce qui concerne la **taille mémoire** utilisée. Par contre, elle nécessite une **étude préliminaire** assez avancée puisque le processus considéré doit être markovien : notamment, si certaines lois d'évolution ne sont pas exponentielles, il faut recourir à l'utilisation d'états fictifs comme expliqué au paragraphe 1.F.

Si on n'est pas limité par la taille mémoire, on peut aussi se ramener à l'étude d'une chaîne de Markov **en observant le processus à des intervalles de temps réguliers**. On retrouve alors les problèmes classiques de la discrétisation du temps. Le délai entre deux observations (le "pas" de la discrétisation) doit être suffisamment court pour que l'on puisse définir aisément une approximation satisfaisante de la loi d'évolution entre deux observations ; par contre, si on observe effectivement pas à pas, la durée de la simulation et l'ordre de grandeur des erreurs d'arrondis sont inversement proportionnels à ce délai.

Du point de vue mathématique, cela revient, pour calculer $\lim_{t \rightarrow \infty} e^{bt}$, à choisir c et $n > 1$ en sorte que c soit une bonne approximation de $e^{b/n}$ puis à calculer, par simulation informatique, $\lim_{m \rightarrow \infty} c^m$: cette remarque évidente

peut permettre de préciser la validité de cette technique ; les "matrices" e^b et c sont des matrices stochastiques (cf. 1.C.3).

Notons aussi que, dans le cas irréductible (c'est à dire le cas normal du point de vue des applications), $e^{b/n}$ a tous ses termes strictement positifs alors que, si b est une matrice creuse (cf. chapitre 5), on choisit, en général, une approximation c de $e^{b/n}$ qui est, elle aussi, une matrice creuse. Nous allons expliciter cette technique sur un exemple très simple.

E.3 File GI/GI/1

Nous avons déjà rencontré des exemples de file unique notamment en 1.E.4, 1.E.5, 1.F.1 et 2.C.2, 3°). La file unique est le module de base dans les systèmes de files d'attente. On considère le cas le plus simple pour lequel il y a une seule classe de clients et un seul serveur. On suppose que la discipline de service est FIFO (First In, First Out), c'est à dire que les clients sont servis dans leur ordre d'arrivées. On suppose que le délai entre deux arrivées suit une loi fixe indépendante de l'état de la file et de son passé : soit X la variable aléatoire associée à ce délai.

De même, on suppose que la durée de service d'un client suit une loi fixe, indépendante de l'état de la file et de son passé : soit Y la variable aléatoire associée à cette durée de service. Dès qu'un service est terminé, le serveur commence le service du client suivant soit aussitôt si ce client est déjà dans la file, soit dès son arrivée.

Soit h le pas de discrétisation choisi, c'est à dire que l'on va étudier le processus aux instants $m h$, pour tous les entiers m , $0 \leq m \leq M$; la durée de la simulation et le temps de calcul sont proportionnels à M ; $M h$ est toujours très grand et correspond à un instant pour lequel le régime stationnaire est (à peu près) atteint depuis longtemps. On "fait comme si" il n'y avait pas de changements d'état entre $m h$ et $(m+1)h$ (quel que soit m) ; par contre, des transitions peuvent avoir lieu aux instants $m h$.

Soit $(x_j)_{1 \leq j < J}$ la séquence croissante de réels positifs telle que, quel que soit j , $1 \leq j < J$, $\text{Proba}[X \leq j h] = x_j$. On va approcher la variable aléatoire X par la variable aléatoire X' discrète dont la loi est définie par $\text{Proba}[X' = J h] = 1 - x_{J-1}$ et, quel que soit j , $1 \leq j < J$, $\text{Proba}[X' = j h] = x_j - x_{j-1}$. On procède de même pour la variable aléatoire Y en considérant la séquence croissante de réels positifs $(y_k)_{1 \leq k < K}$ définie par, quel que soit k , $1 \leq k < K$, $\text{Proba}[Y \leq k h] = y_k$. On pose $x_J := y_K := 1$.

Les séquences (x_j) et (y_k) correspondent aux données statistiques

brutes (étude classique des séries statistiques) : elles ne nécessitent aucun traitement mathématique préalable contrairement à la modélisation par états fictifs (cf. 1.F) ; évidemment, elles doivent être mises en mémoire.

Notons que la simulation simplifiée proposée ici ne convient pas si l'approximation de X par X' n'est pas valable, c'est à dire par exemple, si $E(|X-X'|)$ n'est pas négligeable par rapport à $E(X)$. C'est notamment le cas s'il y a des éventualités dont le rôle est crucial (cas où X est beaucoup plus grand que J h) même si leur probabilité est faible.

Nous verrons plus loin comment en fait, on met en oeuvre la simulation. Nous allons d'abord en donner l'idée de base : pour que l'évolution soit markovienne, il faut que, à chaque pas, c'est à dire à chaque instant m h, l'état prenne en compte le nombre n de clients dans la file, le délai j h qui reste avant l'arrivée du client suivant et, pour $n>0$, la durée k h du service restant à effectuer. On met donc en mémoire ce triplet d'entiers (n, j, k) .

Si on procède pas à pas, à chaque pas on diminue j et k d'une unité si j et k sont strictement positifs : les autres transitions ne peuvent avoir lieu que si $j=0$ ou $k=0$. Supposons, par exemple, que à l'instant $(m-1)h$, l'état e soit $e = (n, j, k)$ avec $j=1$ et $k>1$. Il faut définir l'état $e' = (n', j', k')$ à l'instant m h. On pose évidemment $n' := n+1$ et $k' := k-1$. On utilise un générateur de nombres au hasard $(Z_r)_{r>0}$ tel que chaque variable Z_r suive la loi équidistribuée sur $\{0, 1, \dots, s-1\}$. Supposons que Z_{r-1} soit le nombre précédemment utilisé : j' est alors l'unique entier tel que

$$x_{j'-1} \leq \frac{1}{s} Z_s < x_{j'}$$

On détermine de façon analogue les autres types de transition.

En fait, pratiquement, si J et K sont assez grands et s'il n'y a qu'une seule file (ou, plus généralement, s'il n'y a qu'un petit nombre de files) à considérer, on omet les pas m h pour lesquels on a à la fois $j>1$ et $k>1$. Plus précisément, après chaque transition, on compare j et k ; si, par exemple, on a $j<k$, on pose $n' = n+1$, $k' = k-j$ et on définit j' comme précédemment, ce qui introduit le paragraphe E.4 qui suit. On étudie de même les cas $j>k$ et $j=k$.

La simulation que l'on vient d'étudier admet de nombreuses variantes ou généralisations. Par exemple, si $E(|X-X'|)$ est non négligeable on peut utiliser la séquence $(x_j)_{1 \leq j < 2J}$ telle que, pour $1 \leq j < J$, $\text{Proba}[X \leq j \text{ h}] = x_j$ et, pour $J < j < 2J$, $\text{Proba}[X \leq 2j-J \text{ h}] = x_j$ ou toute autre construction analogue : il faut alors adapter en conséquence les règles de transitions.

On peut aussi étendre ce qui précède au cas d'un système de files

d'attente : on note que la mise en oeuvre de la simulation est assez simple, même si la loi d'évolution du système de files d'attente étudié est complexe. Au contraire, nous verrons par la suite que, dans l'état actuel des connaissances, on ne dispose de formule mathématique simple pour la probabilité stationnaire d'un système de files d'attente que pour certaines lois d'évolution très précises et contraignantes.

Enfin, comme au paragraphe D.5 qui précède, **il faut garder en mémoire...** "tout ce qui nous intéresse" : par exemple, pour chaque entier i , $0 \leq i \leq I$, le nombre de pas m h pour lesquels l'état associé $e = (n, j, k)$ est tel que $n=i$.

E.4 Temps d'arrêt

Dans le paragraphe E.3 qui précède, on a noté que - dans la mesure où cela ne pose pas trop de problèmes techniques - on ne prend en compte que les instants m h où il "se passe quelque chose". La formalisation précise associée à cette remarque revient à étudier le processus à l'aide d'une suite $(u_n)_{n \geq 0}$ de **"temps d'arrêt"**. Même dans le cas où l'ensemble des états est fini, il n'est pas possible de donner une définition rigoureuse de cette expression "temps d'arrêt" sans faire appel à un formalisme mathématique lourd qui sera précisé au chapitre 10. En simplifiant, **un temps d'arrêt est un instant u qui dépend du hasard antérieur à u mais pas du hasard postérieur à u** . Cette définition est valable même si le processus X n'est pas markovien.

Il faut aussi que, pour chaque temps d'arrêt de la suite $(u_m)_{m \geq 0}$, on dispose de suffisamment d'informations pour que l'évolution du processus au delà de u_m ne dépende pas du passé du processus connaissant ces informations : du point de vue théorique on peut considérer que le processus est markovien mais on ne l'observe qu'aux instants $(u_n)_{n \geq 0}$: du point de vue technique, il suffit donc de mettre en mémoire l'information associée à de tels temps d'arrêt.

Reprenons l'exemple de la file GI/GI/1 donné dans le paragraphe précédent. On définit la suite $(u_m)_{m \geq 0}$ d'instants "aléatoires" (les temps d'arrêt) et la suite associée $(v_m)_{m \geq 0}$ d'informations (d'états) de la façon suivante : u_{m+1} est le premier instant postérieur à u_m pour lequel il y a soit un départ, soit une arrivée ; à l'instant u_m , on met en mémoire l'état v_m qui comprend le nombre n de clients dans la file ainsi que l'état v'_m du service et l'état v''_m relativement aux arrivées. Plus précisément, si on

utilise une modélisation par états fictifs (cf. 1.F), en ce qui concerne la loi de service Y , v'_m est un élément de l'ensemble des états fictifs. Au contraire, si on détermine Y par sa fonction de répartition $F(y) := \text{Proba}[Y \leq y]$, du point de vue théorique, l'ensemble des valeurs de v'_m est \mathbb{R}^+ et on prend pour v'_m le délai qui s'est écoulé depuis le début du service : l'évolution ultérieure sera la loi de Y conditionnée par l'évènement $[Y > v'_m]$ (cette expression a un sens théorique très précis) ; du point de vue "technique", comme on l'a fait au paragraphe précédent, dès que le service commence on décide, par un tirage au hasard, de sa durée et, pour chaque temps d'arrêt u_m , on met dans la mémoire v'_m la durée du service restant à effectuer, l'ensemble des valeurs possibles de v'_m étant fini. Notons que le cas $v'_m = 0$ a le même rôle qu'un "état pivot" (cf. 1.F.4).

On procède de même pour l'information v''_m associée aux arrivées. Evidemment on peut utiliser la modélisation par états fictifs pour une loi et la fonction de répartition pour l'autre.

On peut noter que la différence entre les deux méthodes proposées dans ce paragraphe E.4 et dans le paragraphe E.3 qui précède se situe plus au niveau conceptuel que technique.

F. Le progiciel QNAP2

F.1 Langages de simulation

Si un utilisateur souhaite étudier un problème simple et spécifique par simulation, on peut lui conseiller de programmer directement cet exemple. Par contre, dès que l'exemple est complexe, formaliser les diverses transitions est vite fastidieux. Il est alors plus rapide d'utiliser un langage de simulation.

Il y a une très grande variété de tels langages qui se différencient non pas en fonction des méthodes de simulation sous-jacentes mais en fonction des besoins des utilisateurs. Ces divers langages sont commercialisés par des sociétés concurrentes et il n'est pas toujours facile de trouver un livre qui propose une comparaison objective et récente de ces divers langages.

Sans effectuer de comparaison, [Inh] signale les qualités caractéristiques de divers langages, notamment Simone, Simula et May. D'ailleurs, comme cela est noté dans [Ler], le plus souvent "l'utilisateur choisit un langage en fonction de sa disponibilité sur le site d'exploitation,

de sa documentation", etc...

Enfin, et cette évolution est irréversible, comme tous les produits les "langages de simulation" sont soumis à "une économie de marché et non plus à une économie de produit" (A. Riboud) c'est à dire qu'ils sont achetés en fonction des services qu'ils peuvent rendre à l'utilisateur dans un environnement concurrentiel et non pas en fonction de ce qu'ils sont en tant que produit : de ce fait les outils mathématiques proposés ne se limitent plus à la simulation.

Autant que l'auteur ait pu en juger, le progiciel le plus relié aux techniques étudiées dans ce cours est QNAP2 qui est commercialisé par la société Simulog : la suite de ce paragraphe F se propose donc de donner une idée des possibilités des langages de simulation à partir de cet exemple.

F.2 Les déclarations dans QNAP2

QNAP2 est écrit en Fortran ce qui en assure la portabilité. Par contre le "langage QNAP2" est analogue au Pascal (utilisation du parenthésage begin-end, possibilité de créer des "types" en un sens très général, etc...).

Un programme QNAP2 commence par des déclarations où on donne les **noms** des entités utilisées dans le programme : en plus des types classiques (integer, real, boolean, string, ...) on peut notamment utiliser les types suivants : queue, class, class integer, etc...

F.3 Le type queue

Pour chaque queue annoncée dans la déclaration, le programme doit, par la suite, en préciser les caractéristiques : s'il y a plusieurs classes de clients, ces caractéristiques peuvent dépendre de la classe. Ces caractéristiques comportent notamment le nom de la station, la loi de service (exponentielle, hyperexponentielle, Erlang ou coxienne), le "nombre de serveur" (y compris infini ou processor sharing), les routages à la sortie de la station en direction des autres stations (ces routages doivent être fixes et peuvent comporter des changements de classe à la sortie de la station) et la discipline de service (classes prioritaires, fifo, etc...).

F.4 Sémaphore, resource, flag

Lorsque la méthode de résolution est la simulation, QNAP2 permet aussi de prendre en compte les phénomènes de blocage par l'intermédiaire de "stations" sémaphore ou resource, de drapeaux, etc... . A chaque transition, ces stations sont testées pour vérifier que la transition n'est pas

interdite.

Par exemple, une station sémaphore est un compteur qui peut prendre des valeurs entières positives ou négatives : à chaque fois qu'une transition fait appel à ce sémaphore, la transition est annulée si le compteur est négatif ou nul : sinon le compteur est décrémenté de 1 ; d'autres transitions peuvent l'incrémenter.

La station ressource correspond à une procédure un peu plus subtile : en simplifiant, elle délivre - à d'autres stations et sur demande - des jetons (qui peuvent dépendre de la classe, du niveau de priorité, etc...) qui sont disponibles en quantité limitée et qui sont rendus à la fin du service (dans la station qui a effectuée la requête) : évidemment, s'il n'y a pas de jeton disponible, le service est bloqué.

On peut aussi créer des drapeaux (flag) à deux états (set ou unset) ce qui permet de synchroniser des services : évidemment, un service qui nécessite un drapeau est bloqué quand ce drapeau est dans l'état unset.

Bien entendu, ces paragraphes F.3 et F.4 n'épuisent pas les possibilités "constructives" de QNAP2.

F.5 Résolution

La méthode de résolution peut être choisie par l'utilisateur. En fait, les possibilités de résolution actuellement implémentées sur QNAP2, tout en étant d'un haut niveau mathématique, sont très loin d'exploiter la totalité des techniques proposées dans ce cours. Notamment, dès qu'il y a des phénomènes de blocage, la solution est automatiquement calculée par simulation.

Quand la résolution est effectuée par simulation, le générateur de nombres au hasard sous-jacent est changé régulièrement : quand c'est possible, ce changement est effectué "aux points de régénération" (cf. 8.E). On obtient ainsi une famille de "réalisations du hasard" et QNAP2 donne les intervalles de confiance associés à cet ensemble de simulation.

F.6 Autres aspects

Evidemment - et ceci constitue un gain de temps considérable pour l'utilisateur - QNAP2 propose les "commodités" que l'on est en droit d'attendre d'un langage de simulation de haut niveau.

Les paramètres peuvent être donnés en mode interactif. L'utilisateur peut obtenir en sortie, par des commandes très simples, les caractéristiques classiques telles que temps de traitement, temps d'occupation, lois marginales, etc... Il peut aussi obtenir les informations spécifiques de son choix.

Plusieurs visualisations des résultats sont proposées : tableaux, courbes, histogrammes, etc...

G. Conclusion

Compte tenu de sa souplesse et de sa commodité d'utilisation, la simulation informatique, comme la marche à pieds, est et restera **un outil de base incontournable**. Tant qu'on se restreint à sa mise en oeuvre, elle nécessite peu de connaissances théoriques. Toutefois, il peut ne pas être inutile de passer quelques instants à choisir une bonne paire de chaussures, ce choix n'étant pas forcément le même selon que l'on projette d'errer en Sologne ou aux abords du Mont Blanc. Par exemple, il est peut-être dommage de se priver de la modélisation par états fictifs si celle-ci permet un gain considérable en temps de calculs.

En ce qui concerne la validité de la simulation, même en régime homogène stationnaire, il n'y a pas, actuellement, de résultats généraux satisfaisants au niveau des applications. Par contre, diverses études théoriques offrent des éclairages révélateurs. Par exemple, les occurrences d'états dans un "voisinage" donné sont d'autant plus régulières que le coefficient de contraction (cf. 2.A) est faible. A contrario, si ce coefficient de contraction - des matrices a^k où a est la matrice de transition sous-jacente à la simulation proposée - est très proche de 1, c'est à dire s'il y a un "mauvais" mélange et **si certains états "peu accessibles" ont un rôle crucial** (cas de la centrale nucléaire qui explose), toute étude par simulation risque de conduire à des conclusions erronées.

De même, et contrairement à une idée encore répandue, prolonger indéfiniment une simulation **n'augmente pas sa fiabilité** à partir d'un certain seuil (cf. [Bec]).

L'apport de la modélisation mathématique ne se limite pas à ces coups de projecteurs. Nous verrons, au chapitre 4 que, sous certaines hypothèses - précises mais assez souvent satisfaites - la solution stationnaire a une forme connue explicite facile à exploiter, alors que, eu égard au nombre d'états considérés, toute étude par simulation est presque inconcevable - (sauf si on tient à aller à pieds à Jérusalem pour des raisons idéologiques).

De même, nous verrons au chapitre 5 des exemples pour lesquels on dispose d'algorithmes de calculs des probabilités stationnaires nettement plus performants que ceux associés à la simulation.

Par ailleurs, en général, l'étude mathématique, exacte ou approchée, se prête beaucoup mieux aux études de sensibilité par rapport aux paramètres.

Enfin, le 9 octobre 1890, fallait-il préjuger de l'avenir de l'aviation au vu de l'envolée de l'éole sur une cinquantaine de mètres : les résultats mathématiques aujourd'hui effectivement utilisables sont très limités mais porteurs d'espoir d'améliorations substantielles.

Quoiqu'il en soit, pour les études coûteuses et complexes, l'avenir privilégiera probablement les **logiciels de haut niveau** (les "progiciels" : cf. 3.F). Ces logiciels conviviaux permettront à l'utilisateur d'exposer son problème ; la méthode de résolution, avec ou non une part de simulation, sera décidée par le concepteur du logiciel en fonction du problème posé. Pour autant, cela ne signifie pas que l'utilisateur aura intérêt à méconnaître le fonctionnement "interne" des logiciels qu'il utilise.

Or, "s'il est certain que des outils logiciels d'environnement de programmation et d'intelligence artificielle seront développés, par les informaticiens, pour faciliter l'utilisation pratique des ordinateurs, il est non moins certain que les méthodes numériques adaptées aux différents domaines du calcul scientifique ne seront pas inventées par les calculateurs" (P. Leca et F.X Roux, Matapli, oct. 1988).

Actuellement, le plus souvent, la priorité est encore donnée aux facilités d'utilisation mais les progrès dans ce domaine relativement à un type précis de problème ont un horizon fini alors que les améliorations potentielles des méthodes numériques sont illimitées : ces améliorations - qui commandent quant au fond les services que l'on peut proposer à l'usager - seront donc, à l'avenir, un argument essentiel du "marketing".

Chapitre 4

Solution à forme produit**A Serveur central****A.1 Réseau de files d'attente**

Lors de la première partie de ce cours - partie constituée des trois premiers chapitres - on a présenté le cadre général associé aux processus markoviens dont l'ensemble des états est "presque" fini. Dans la deuxième partie, nous allons donner les résultats qui nous semblent les plus importants au niveau des applications et qui permettent de calculer, de façon théoriquement **exacte**, les probabilités stationnaires (cf. 2.D) de certains systèmes markoviens.

Nous commençons cette deuxième partie par un chapitre centré sur les réseaux de files d'attente : ce n'est pas que cette notion soit plus fréquente qu'une autre au niveau des applications. Le système d'hypothèses étudié dans ce chapitre est important parce que c'est un système relativement général pour lequel **la solution stationnaire a une forme explicite exploitable**. Comme on l'a déjà noté, il serait peu judicieux d'étudier par simulation un phénomène que l'on peut modéliser par un système markovien qui satisfait (presque) aux hypothèses données dans ce chapitre.

Ces hypothèses seront précisées au cours du chapitre. Le **langage** utilisé sera le langage usuel des réseaux de files d'attente, c'est à dire que l'on étudie des **clients** qui passent dans diverses **stations** pour y être **servis**. Il peut y avoir plusieurs classes de clients et plusieurs types de services et de disciplines de service.

Bien entendu, il ne s'agit que d'un langage. Les clients peuvent être des programmes qu'on étudie à un certain niveau d'un ordinateur, des appels à un standard téléphonique ou des messages qui traversent un réseau de transmission, etc... Une station de service peut être associée à un "bus" de liaison dans un système électronique, au fait, pour des véhicules, de tomber en panne ou à l'ensemble des clients qui veulent téléphoner et en sont à leur troisième tentative pour obtenir la ligne, etc...

A.2 Notations générales

Les notations générales sont celles de l'ensemble de ce cours (cf., notamment, 2.D) ; plus précisément : E est l'ensemble des états, a est une

fonction positive définie sur $(E \times E)$ et on cherche une famille $(q(u))_{u \in E}$ de nombres positifs telle que les deux propriétés suivantes soient satisfaites :

$$(4A1) \quad q(u) \sum_{v \in E} a(u,v) = \sum_{v \in E} q(v) a(v,u)$$

$$(4A2) \quad \sum_{u \in E} q(u) = 1$$

Pour tout élément u de E , on suppose que $a(u,u) = 0$ (on peut toujours se ramener à ce cas). En général la famille q est unique. Cette famille q peut être considérée comme la **probabilité stationnaire** du processus markovien $(X_t)_{t \in \mathbb{R}^+}$ dont la loi d'évolution est définie, pour $u \neq v$, par :

$$a(u,v) := \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba}[X_{t+h} = v \mid X_t = u] \right\}$$

Dans ce qui suit, E sera souvent une partie de \mathbb{N}^H , H étant un ensemble fini. Pour tout élément $u := (u_h)_{h \in H}$ de \mathbb{N}^H on définit $(u+e_k)$ comme suit : $(u+e_k)$ est l'élément $v := (v_h)_{h \in H}$ de \mathbb{N}^H défini par :

$$v_k := 1+u_k \text{ et, pour } h \neq k, v_h := u_h.$$

Par exemple, si H est l'ensemble des stations et u_h le nombre de clients à la station h (ce qui définit l'état u), alors $(u+e_k)$ est l'état u sauf qu'il y a **un client de plus dans la station k** .

On définit de même $u-e_k$, $u-e_j-e_k$, etc... Suivant les conventions usuelles, $u(h) = u_h$, etc...

A.3 Hypothèses du serveur central

1°) Soit H un ensemble fini. On suppose que l'ensemble E des états est contenu dans \mathbb{N}^H ; on suppose aussi que cet ensemble E est tel que, si (u,h) appartient à $(E \times H)$ et si $u_h > 0$, alors $(u-e_h)$ **appartient aussi à E** . Cet ensemble E peut être fini (réseau fermé) ou infini (réseau ouvert).

2°) Soit s une fonction positive définie sur $(H \times \mathbb{N})$; on suppose que, pour tout élément h de H , $s(h,0) = 0$ et, pour $i > 0$, $s(h,i) > 0$.

3°) Soit r une fonction positive définie sur $(H \times E)$ qui satisfait aux propriétés suivantes (quels que soient h et k éléments de H) :

$$(4A3) \quad (u + e_h) \notin E \text{ implique } r(h, u) = 0$$

$$(4A4) \quad (u + e_h) \in E \text{ implique } r(h, u) > 0$$

$$(4A5) \quad \text{pour tout élément } u \text{ de } E, \\ r(h, u - e_h) r(k, u - e_h - e_k) = r(k, u - e_k) r(h, u - e_h - e_k)$$

Des exemples de telles fonctions r seront données plus loin. L'adéquation de la relation 4A5 sera expliquée au paragraphe A.4 qui suit.

4°) La fonction a (cf. A.2) est définie par : quel que soit (h, u) élément de $(H \times E)$,

$$a(u, u + e_h) := r(h, u) \quad \text{et} \quad a(u, u - e_h) := s(h, u_h)$$

Dans tous les autres cas, $a(u, v) := 0$.

5°) Interprétations

Une première interprétation consiste à considérer qu'il y a **une seule classe de clients ; H est alors l'ensemble des stations "internes"** ; il y a aussi une station "externe". L'état u est caractérisée par le nombre de clients u_h dans chaque station h ; $s(h, i)$ est le **taux de service** à la station h quand il y a i clients dans cette station ; un client qui est servi à la station interne h va à la station externe ; $r(h, u)$ est le **taux de service à la station externe multiplié par la probabilité pour que le client servi aille à la station h quand l'état est u .**

En fait, de même que dans les exemples qui seront donnés par la suite, on peut aussi supposer que chaque élément h de H est un couple $h := (h', h'')$ où h' est une classe de clients et h'' est une station. La présentation qui précède signifie que le taux de service des clients de classe h' à une station h'' ne dépend que du nombre de clients de cette classe h' qui sont dans la station h'' . Cette restriction est uniquement pour alléger la présentation de ce premier paragraphe (cf. la remarque A.5.3°) et le paragraphe D qui suivent).

A.4 Probabilité stationnaire

Proposition : Soit o l'élément u de E défini par $u_h = 0$ pour tout élément h de H .

1°) Il existe une fonction unique t définie sur E telle que $t(o) = 1$ et, quel que soit (u, h) élément de $(E \times H)$,

$$(4A6) \quad t(u) = t(u - e_h) r(h, u - e_h)$$

2°) Pour tout élément (h, u) de $(H \times E)$, on pose :

$$(4A7) \quad s'(h, i) := \prod_{j \leq i} s(h, j) \quad \text{et}$$

$$(4A8) \quad q'(u) := t(u) / \left\{ \prod_{h \in H} s'(h, u_h) \right\}$$

On suppose que $\sum_{u \in E} q'(u) = c < +\infty$. On pose $q(u) := q'(u)/c$:

la famille $(q(u))_{u \in E}$ est la probabilité stationnaire.

Preuve :

1°) Soit u un élément de E ; soit $(u_j)_{1 \leq j \leq m}$ une séquence d'éléments de E telle que $u_1 = 0$, $u_m = u$ et, quel que soit j , $1 \leq j < m$, il existe $h(j)$ élément de H tel que $u_{j+1} = u_j + e_{h(j)}$; alors on peut poser :

$$(4A9) \quad t(u) := \prod_{j=1}^{m-1} r(h(j), u_j)$$

En effet, cette quantité $t(u)$ ne dépend que de u et non pas de la séquence $(u_j)_{1 \leq j \leq m}$, compte tenu de la relation 4A5 ; cette fonction t est la seule qui satisfasse à 4A6 et qui soit telle que $t(0) = 1$.

2°) La "matrice" a est évidemment irréductible. On pose, pour tout élément (u, h) de $(E \times H)$:

$$\begin{aligned} \tau(u, h) &:= q(u) s(h, u_h) \\ \tau(u, h) &:= q(u - e_h) r(h, u - e_h) \\ \tau(u, h) &:= q(u) r(h, u) \\ \tau(u, h) &:= q(u + e_h) s(h, 1 + u_h) \end{aligned}$$

En langage imagé, $\tau(u, h)$ est le "taux de probabilité" de "**quitter**" l'état u , un client "**quittant**" la station h ; $\tau(u, h)$ est le "taux de probabilité" d'"**atteindre**" l'état u , un client "**rentrant**" dans la station h . Les autres $\tau(.,.)$ ont une interprétation analogue. La définition de q donne

immédiatement que, pour tout élément (u, h) de $(E \times H)$, on a :

$$(4A10) \quad \tau(u, h)^\uparrow = \tau(u, h)^\downarrow \text{ et }$$

$$(4A11) \quad \tau(u, h)^\uparrow = \tau(u, h)^\downarrow$$

(si $(u + e_h)$ n'appartient pas à E , les deux termes de 4A11 sont nuls).

On pose :

$$\tau(u)^\uparrow := \sum_{v \in E} q(u) a(u, v) \text{ et } \tau(u)^\downarrow := \sum_{v \in E} q(v) a(v, u)$$

$\tau(u)^\uparrow$ est le taux de quitter l'état et $\tau(u)^\downarrow$ est le taux d'atteindre l'état u .

On a :

$$\tau(u)^\uparrow = \sum_{h \in H} [\tau(u, h)^\uparrow + \tau(u, h)^\downarrow]$$

et

$$\tau(u)^\downarrow = \sum_{h \in H} [\tau(u, h)^\downarrow + \tau(u, h)^\uparrow]$$

Les relations 4A10 et 4A11 impliquent donc que $\tau(u)^\uparrow = \tau(u)^\downarrow$; la famille $(q(u))_{u \in E}$ satisfait donc aux conditions d'équilibre ; on a choisi c en sorte que $\sum_{u \in E} q(u) = 1$ donc q est la probabilité stationnaire.

A.5 Remarques

1°) La preuve qui précède est typique de la plupart des preuves que l'on rencontrera dans ce chapitre. D'une part, on aura très souvent à faire intervenir des quantités telles que $\tau(u, h)^\uparrow$ (par exemple). D'autre part, on prouvera que q est la probabilité stationnaire en montrant que q **satisfait aux conditions d'équilibre** et en utilisant **l'unicité** de cette solution, à une constante multiplicative près. Enfin, le plus souvent, on prouvera une famille de relations plus fortes que les conditions d'équilibre.

2°) La propriété 4A10 est appelée **la balance locale** (pour la station h) : elle est assez souvent satisfaite. La propriété 4A11 est rarement satisfaite (cf. les réseaux réversibles au chapitre 9).

3°) Soit K une partie de H . On pose :

$$\tau(u, K)^\uparrow := \sum_{h \in K} \tau(u, h)^\uparrow \text{ et } \tau(u, K)^\downarrow := \sum_{h \in K} \tau(u, h)^\downarrow$$

On a évidemment

$$\tau(\overset{\uparrow}{u}, \overset{\uparrow}{K}) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{K})$$

On peut donc modifier les taux de service comme indiqué au paragraphe D qui suit.

A.6 Routages fixes

Nous allons maintenant donner des exemples de fonctions r satisfaisant à la relation 4A5 et de "décors" associés. Le cas le plus simple est le cas où il y a une seule classe de clients et **les "routages" ne dépendent pas de l'état**. Plus précisément, soit $(x_h)_{h \in H}$ une famille de réels strictement positifs telle que $\sum_{h \in H} x_h = 1$; x_h est la probabilité (le "routage"), pour un client qui quitte la station externe, d'être dirigé vers la station h .

Pour tout élément u de E , posons $u' := \sum_{h \in H} u_h$; soit d une fonction positive définie sur \mathbb{N} telle que $i > 0$ et $d(i) > 0$ impliquent $d(i-1) > 0$; $d(i)$ est le taux de service dans la station externe quand il y a i clients dans l'ensemble des stations internes. Plus précisément, on définit r par, quel que soit (h, u) élément de $(H \times E)$:

$$r(h, u) := d(u') x_h$$

La vérification de la propriété 4A5 est alors immédiate.

Pour tout entier i , posons :

$$d'(i) := \prod_{j < i} d(j)$$

La probabilité stationnaire vaut :

$$(4A12) \quad q(u) = \left\{ \prod_{h \in H} [(x_h)^{u(h)} / s'(h, u_h)] \right\} d'(u') / c$$

où c est la "constante de normalisation".

Rappelons que, si l'ensemble E est infini, ce qui précède n'a de sens que si $\sum_{u \in E} q(u)c < +\infty$.

Ce système d'hypothèses (serveur central avec des routages fixes) est souvent utilisé, en première approximation, **pour modéliser le fonctionnement d'un ordinateur** : la station externe correspond à l'"unité centrale" de l'ordinateur ; les stations "internes" sont les "périphériques"

(imprimantes, disques, etc...), ces périphériques étant en général classées suivant la vitesse avec laquelle la mémoire peut y accéder. Le réseau est fermé.

En général, la mémoire centrale fonctionne en temps partagé : d ne dépend pas du nombre de clients dans l'unité centrale. Au contraire, en général, chaque périphérique est modélisée par une station avec w serveurs (on a souvent $w = 1$) : le taux de service de la station h associé est alors $s(h,i) := w' \cdot \inf(w,i)$ où w' est le taux de service de chaque serveur.

Dans ce qui précède, on a considéré le cas d'une seule classe de clients ; on peut aussi supposer que chaque élément h de H est un couple (h', h'') où h' est une classe et h'' une station ; quand un client est servi à la station externe, la probabilité qu'il devienne de classe h' et qu'il aille à la station h'' vaut x_h .

A.7 Routages variables

On se limite à une seule classe de clients.

Rappelons que, en A.3, la seule hypothèse portant sur l'ensemble E est que (h,u) élément de $(H \times E)$ et $u_h > 0$ implique que $(u - e_h)$ appartient à E . Si on s'autorise - en première approximation - de choisir des ensembles E finis et "surprenants", les conditions 4A3, 4A4 et 4A5 admettent une infinité de solutions et celles-ci sont suffisamment souples pour permettre d'approcher un grand nombre de situations concrètes.

Nous allons donner un exemple pour lequel r peut être exprimé sous forme d'une formule simple. On définit d et d' comme en A.6.

Soit z réel. Soit $(w_h)_{h \in H}$ une famille de réels positifs. Pour tout élément u de E , on pose :

$$u'' := \sum_{h \in H} (w_h + z u_h)$$

et, pour tout élément h de H :

$$(4A13) \quad x_h(u) := (w_h + z u_h) / u'' \quad \text{et, pour } (u + e_h) \in E$$

$$(4A14) \quad r(h,u) := d(u') x_h(u)$$

Si $(u + e_h)$ n'appartient pas à E , on pose $r(h,u) := 0$.

Si on fait $z = 0$, on retrouve exactement les hypothèses données en A.6 (avec $x_h = w_h$) mais, évidemment, la condition 4A13 est beaucoup plus

générale que celle considérée en A.6.

On suppose que, pour tout élément u de E , $x_h(u) \geq 0$ (cette hypothèse est indispensable ; en général, on choisit $z < 0$).

On vérifie alors immédiatement que la fonction r définie par 4A14 satisfait à la propriété 4A5.

Pour fixer les idées, considérons un "décor" associé à ce système d'hypothèses (cf. [Pel.2], [Pel.4] ou [Pel.6]). On prend, par exemple, $z = -1$. La station centrale modélise l'ensemble des véhicules d'une entreprise nationale qui sont en état de marche. Dans chaque département h , cette entreprise dispose d'un centre de maintenance dans lequel il y a w_h places. E est défini par : u appartient à E si et seulement si, quel que soit le département h , $u_h \leq w_h$. Le taux de "service" de la station centrale peut être choisi diversement : on peut supposer qu'il est fixe (on surcharge les véhicules en état de marche) ou de la forme $(n-u')^b$ où n est le nombre total de véhicules. Quand un véhicule tombe en panne, le hasard de sa position sur le sol français et des procédures de routages adéquates font que sa probabilité d'être envoyé vers le centre de maintenance h est proportionnelle au quotient de la place qui reste dans ce centre par la place totale restant disponible.

On note que cette loi de "dispatching" est "**hiérarchique**" au sens suivant : si on regroupe les départements en régions, tout se passe comme si on dirigeait les véhicules à réparer vers les régions proportionnellement au nombre de places disponibles puis, à l'intérieur de chaque région, vers chaque département toujours suivant la même règle.

Redisons que ceci **est un exemple** pour montrer que la condition 4A5 est réaliste mais on peut utiliser des routages beaucoup plus généraux. Rappelons aussi qu'il est exceptionnel que la probabilité stationnaire puisse s'écrire suivant une formule explicite simple (comme cela est le cas dans ce chapitre) : pour disposer d'une telle formule, il ne faut donc pas hésiter à choisir une modélisation imparfaite, éventuellement avec un ensemble E d'états "surprenant" (cf. le début de ce paragraphe A.7).

Notons aussi qu'une étude systématique de la balance locale pour un réseau comportant trois stations a été effectué dans [L1b] et [L1P].

B. Routages fixes, monoclasse

B.1 Hypothèses : cas fermé

Soit H un ensemble fini : H est l'ensemble des stations (il n'y a pas d'autres stations). Soit n un entier, $n > 0$. L'ensemble E des états est une partie de \mathbb{N}^H : plus précisément, $u := (u_h)_{h \in H}$ appartient à E si et seulement si $\sum_{h \in H} u_h = n$; u_h est le nombre de clients dans la station h . Il y a une seule classe de clients.

Pour chaque station h , on se donne $s(h,i)$ le taux de service quand il y a i clients dans cette station (comme en A.3). On suppose $s(h,0) = 0$ et, pour $i > 0$, $s(h,i) > 0$.

Par ailleurs, soit r une fonction positive définie sur $(H \times H)$. On suppose que la "matrice" associée à r est **stochastique** (cf. 1.B.3) et **irréductible** (cf. 2.B.3) ; autrement dit, d'une part, pour tout élément i de H , on a :

$$\sum_{j \in H} r_{i,j} = 1$$

d'autre part (cf. 2.D.3), il existe une et une seule famille $(x_i)_{i \in H}$ de nombres positifs telle que $\sum_{i \in H} x_i = 1$ et, pour tout élément i de H :

$$(4B1) \quad x_i = \sum_{j \in H} x_j r_{j,i}$$

On supposera, ce qui n'est pas une restriction, que, pour tout élément i de H , $r_{i,i} = 0$.

La fonction r définit les routages : plus précisément, un client qui est servi à la station i va dans la station j avec la probabilité $r_{i,j} := r(i,j)$. Autrement dit, la matrice a est définie par :

$$a(u, u + e_j - e_i) := s(i, u_i) r_{i,j}$$

et $a(u,v) = 0$ s'il n'existe pas de couple (i,j) tel que $v = u + e_j - e_i$.

La matrice a est irréductible (puisque r l'est). Les hypothèses données dans ce paragraphe B.1 généralisent celles proposées en A.6.

B.2 Théorème de Jackson

Soit q' la fonction définie sur E par :

$$(4B2) \quad q'(u) := \prod_{h \in H} \{ x_h^{u(h)} / s'(h, u_h) \}$$

Rappelons que $u_h := u(h)$ et $s'(h,i) := \prod_{j \leq i} s(h,j)$.

On pose $c := \sum_{u \in E} q'(u)$ et

$$q(u) := q'(u)/c$$

Alors q est la probabilité stationnaire (cf. [Jac]) associée aux hypothèses données en B.1. (cf. aussi [Buz], [GoN], etc...).

Preuve :

Pour tout élément (h,u) de $(H \times E)$ on pose (cf. la preuve de A.4) :

$$\begin{aligned} \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) &:= q(u) s(h, u_h) \\ \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h}) &:= \sum_{j \neq h} q(u + e_j - e_h) s(j, 1 + u_j) r_{j,h} \end{aligned}$$

La définition de q' implique (cf. 4B2) :

$$\begin{aligned} \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h}) &= \sum_{j \neq h} q(u) x_j s(h, u_h) r_{j,h} / x_h \\ &= q(u) s(h, u_h) \sum_{j \neq h} x_j r_{j,h} / x_h \end{aligned}$$

La relation 4B1 implique alors :

$$(4B3) \quad \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h}) = q(u) s(h, u_h) = \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h})$$

Or, on a aussi

$$\tau(\overset{\uparrow}{u}) = \sum_{h \in H} \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) \quad \text{et} \quad \tau(\overset{\downarrow}{u}) = \sum_{h \in H} \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h})$$

Les relations 4B3 impliquent donc $\tau(\overset{\uparrow}{u}) = \tau(\overset{\downarrow}{u})$; q est donc bien la probabilité stationnaire (unicité).

En fait, comme en A.6, on a prouvé des relations plus fortes que les relations d'équilibre : pour chaque station h , il y a "**balance locale**" dans cette station (relation 4B3). Par contre, en général, $\tau(\overset{\uparrow}{u}, \overset{\downarrow}{h}) \neq \tau(\overset{\downarrow}{u}, \overset{\uparrow}{h})$.

B.3 Cas ouvert

C'est le cas où, en plus des stations considérées précédemment, il y a une "station externe" dans laquelle il y a une "infinité de clients".

L'ensemble E est alors \mathbb{N}^H tout entier ; un élément $u := (u_h)_{h \in H}$ de E est donc encore caractérisé par le nombre u_h de clients dans chaque station "interne" h .

On pose $H' := H \cup \{0\}$ et on suppose maintenant que la fonction r qui définit les "routages" est une fonction positive définie sur $(H' \times H')$. On suppose que la "matrice" associée à r est stochastique et irréductible c'est à dire (cf. 2.D.3) que, d'une part, pour tout élément i de H' , on a

$$\sum_{j \in H'} r_{i,j} = 1$$

d'autre part, il existe une et une seule famille $(x_i)_{i \in H'}$ de nombres positifs telle que $x_0 = 1$ et, pour tout élément i de H' ,

$$x_i = \sum_{j \in H'} x_j r_{i,j}$$

On suppose que, pour tout élément i de H' , $r_{i,i} = 0$. La fonction r définit les routages, comme en B.1.

Pour tout élément u de E, on pose, comme en A.6, $u' := \sum_{h \in H} u_h$; soit (comme en A.6), une fonction strictement positive d définie sur \mathbb{N} : $d(k)$ est le taux de service dans la station externe quand le nombre de clients

dans l'ensemble des stations internes est k ; on pose $d'(i) = \prod_{j < i} d(j)$.

La fonction a qui régit la loi d'évolution est définie par, quel que soit $(i,j,u) \in (H \times H \times E)$:

$$a(u, u+e_j-e_i) := s(i, u_i) r_{i,j}$$

$$a(u, u+e_i) := d(u') r_{0,i}$$

$$a(u, u-e_i) := s(i, u_i-1) r_{i,0}$$

$a(u,v) = 0$ s'il n'existe pas d'indices i (et j) tels que $v = u+e_j-e_i$ ou $v = u-e_i$.

On pose alors :

$$q'(u) := d'(u') \prod_{h \in H} \{ x_h^{u(h)} / s'(h, u_h) \}$$

On suppose que

$$c := \sum_{u \in H} q'(u) < +\infty \quad (\text{condition de non-explosion})$$

La probabilité stationnaire est alors la famille q définie par, pour tout élément u de E , $q(u) := q'(u)/c$.

La preuve est tout à fait analogue à celle proposée en B.2. Plus précisément, pour tout élément h de H on vérifie facilement que :

$$\tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h})$$

Il y a donc balance locale pour chaque station, y compris la station externe.

B.4 La constante de normalisation

L'intérêt essentiel des hypothèses considérées dans ce chapitre est, évidemment, dû au fait que la probabilité stationnaire admet une forme explicite simple. Ceci permet de considérer le cas où il y a un très grand nombre d'états ce qui serait impossible si, toutes choses égales par ailleurs, on procédait par simulation.

Or, quand le nombre d'états est très grand, il n'est pas toujours techniquement facile de calculer $c := \sum_{u \in E} q'(u)$: il y a donc de nombreux articles donnant des algorithmes de calcul de cette constante de normalisation c : cf. notamment, [BhW], [BrB] et [Buz].

Ces algorithmes portent sur le cas monoclasse (comme dans ce paragraphe) ou multiclassé (cf. les paragraphes suivants). Il n'est évidemment pas possible de résumer ici ces divers algorithmes. Nous allons seulement en donner une idée de base à partir d'un exemple.

On suppose que E est l'ensemble des éléments u de \mathbb{N}^H tels que

$$\sum_{h \in H} u_h = n \quad (\text{cas fermé})$$

On suppose que q' admet une forme produit c'est à dire qu'on suppose qu'il existe une fonction b définie sur $(H \times \mathbb{N})$ telle que, pour tout élément u de E , on a :

$$(4B4) \quad q'(u) = \prod_{h \in H} b_{h, u(h)}$$

Soit $m := \text{card}(H)$. A une modification de notations près, on peut toujours supposer que l'ensemble fini H est l'ensemble $H := \{1, \dots, m\}$.

Pour tout élément h de H , on pose :

$$(4B5) \quad \begin{aligned} g_h(z) &:= \sum_{j=0}^n b_{h,j} z^j \quad \text{et} \\ f_h(z) &:= \prod_{j \leq h} g_j(z) \end{aligned}$$

Soit w la fonction définie sur $(H \times \mathbb{N})$ par :

$$(4B6) \quad f_h(z) = \sum_{j=0}^n w_{h,j} z^j$$

Il résulte de 4B4 et des propriétés formelles des "fonctions génératrices" que $c := \sum_{u \in E} q'(u)$ est le coefficient de z^m dans le développement "en série entière" 4B6 de $f_m(z)$.

L'"astuce" consiste alors tout simplement à remarquer qu'on **diminue considérablement le nombre de calculs à effectuer** si on procède de la façon suivante : au lieu de calculer toutes les valeurs possibles de $q'(u)$ et d'en faire la somme, on calcule les coefficients $w_{h,j}$ en raisonnant par récurrence croissante sur h ; plus précisément, la relation 4B5 implique :

$$w_{h,j} = \sum_{k=0}^j b_{h,k} w_{h-1,j-k}$$

Cette technique de calcul de la constante de normalisation permet aussi d'obtenir rapidement d'autres valeurs importantes telles que, par exemple, la probabilité pour une file d'être vide.

B.5 Sensibilité

Ce qui suit généralise [Lin] (cf. le paragraphe C qui suit). On suppose que l'ensemble E des états est contenu dans \mathbb{N}^H . Considérons d'abord le cas où cet ensemble E est fini.

On suppose qu'on a les propriétés suivantes :

a) La probabilité stationnaire associée à l'état u appartenant à E vaut $q(u) := c q'(u)$ avec

$$(1/c) = \sum_{u \in E} q'(u)$$

b) Il existe une fonction positive b définie sur $(H \times E)$ telle que, pour tout élément u de E , on a :

$$q'(u) = \prod_{h \in H} b(h,u)$$

c) La fonction b dépend d'un paramètre s réel et il existe une fonction g définie sur $(H \times E)$ telle que, pour tout élément u de E , on a :

$$\frac{\partial}{\partial s} [b(h,u)] = b(h,u) g(h,u)$$

d) f est une fonction réelle définie sur E .

Soit X la "variable aléatoire" associée à l'état du processus en régime stationnaire, c'est à dire que, pour tout élément u de E ,

$$\text{Proba } [X = u] = q(u)$$

On a alors, (avec les conventions usuelles associées à l'étude des variables aléatoires) :

$$\frac{\partial}{\partial s} E[f(X)] = E \left[\frac{\partial f}{\partial s}(X) \right] + \sum_{h \in H} \text{cov} [f(X), g(h,X)]$$

Vérification

On a (propriété classique) :

$$E(f(X)) = c \sum_{u \in E} f(u) \prod_{h \in H} b(h,u)$$

donc

$$\frac{\partial}{\partial s} [E(f(X))] = \alpha + \beta + \gamma \quad \text{avec}$$

$$\alpha = \frac{\partial c}{\partial s} \sum_{u \in E} f(u) \prod_{h \in H} b(h,u)$$

$$\beta = c \sum_{u \in E} \frac{\partial}{\partial s} f(u) \prod_{h \in H} b(h,u)$$

$$\beta = E \left[\frac{\partial f}{\partial s}(X) \right]$$

$$\gamma = c \sum_{u \in E} f(u) \sum_{h \in H} \left\{ \frac{\partial}{\partial s} [b(h,u)] \prod_{k \neq h} b(k,u) \right\}$$

$$= c \sum_{u \in E} f(u) \sum_{h \in H} \left\{ g(h,u) \prod_{k \in H} b(k,u) \right\}$$

$$= \sum_{h \in H} \sum_{u \in E} f(u) g(h,u) q(u)$$

$$\gamma = \sum_{h \in H} E[f(X) g(h,X)]$$

Par ailleurs,

$$\begin{aligned}
 \frac{\partial c}{\partial s} &= -c^2 \sum_{u \in E} \frac{\partial}{\partial s} [q'(u)] \\
 &= -c^2 \sum_{u \in E} \sum_{h \in H} \frac{\partial}{\partial s} b(h,u) \prod_{k \neq h} b(k,u) \\
 &= -c \sum_{u \in E} \sum_{h \in H} g(h,u) q(u) \\
 \frac{\partial c}{\partial s} &= -c \sum_{h \in H} E[g(h,X)]
 \end{aligned}$$

d'où

$$\begin{aligned}
 \alpha &= - \sum_{h \in H} E[g(h,X)] \sum_{u \in E} f(u) q(u) \\
 \alpha &= - \sum_{h \in H} E[g(h,X)] E[f(X)]
 \end{aligned}$$

ce qui donne

$$\alpha + \gamma = \sum_{h \in H} \text{cov}[f(X), g(h,X)]$$

ce qui implique le résultat annoncé.

Evidemment, si E n'est pas fini, tout ce qui précède reste valable dans la mesure où toutes les familles considérées sont sommables.

C. Routages fixes, multiclasse

C.1 Hypothèses : cas fermé

1°) Les hypothèses (cf. [BCMP]) que nous allons introduire maintenant sont plus générales que celles données en B.1 : c'est uniquement pour la commodité du lecteur et au vu de l'importance du cadre introduit en B.1 que nous avons traité d'abord le cas monoclasse.

2°) Nous considérons donc maintenant le cas où il y a **plusieurs classes de clients mais les routages restent fixes**. Dans ce paragraphe C, nous supposons qu'un état peut être caractérisé par le nombre de clients de chaque classe dans chaque station ; plus précisément, si on définit ainsi l'ensemble E des états, on suppose que, relativement à cet ensemble E d'états, l'évolution est markovienne : on dit parfois que E est constitué de "**macro-états**". Si, dans certaines stations, les disciplines de service tiennent compte de l'ordre d'arrivée des clients, cet ordre doit apparaître au

niveau des éléments de E ; on dit parfois que l'on considère des "**micro-états**" : ce cas sera traité au paragraphe F qui suit.

3°) la modélisation usuelle consiste à introduire un ensemble de stations et un ensemble de classes. Pour faciliter la lecture et l'écriture, il est important d'alléger au maximum les notations. Nous allons donc nous donner un seul ensemble H , chaque élément h de H pouvant être un couple (h', h'') où h' est une classe de clients et h'' une station ; d'autres interprétations peuvent être envisagées. Les routages r_{ij} sont définis exactement comme en B.1 (on suppose que la "matrice" r est stochastique et irréductible et que la "diagonale principale" de r est nulle). La famille $(x_i)_{i \in H}$ est définie comme en B.1 (relations 4B1). Si i et j sont deux couples (i', i'') et (j', j'') où i' et j' sont des classes de clients et i'' et j'' sont des stations, l'interprétation des routages revient à dire qu'un client de la classe i' qui vient d'être servi à la station i'' a la probabilité r_{ij} de rentrer dans la station j'' en étant devenu un client de la classe j' ; on peut avoir $i' = j'$ ou $i'' = j''$. L'ensemble E des états est contenu dans \mathbb{N}^H ; un élément $u := (u_h)_{h \in H}$ de E est donc caractérisé par la famille des nombres u_h : par exemple, si $h = (h', h'')$ où h' est une classe de clients et h'' une station, u_h est le nombre de clients de la classe h' dans la station h'' . Soit n le nombre total de clients. L'ensemble E est contenu dans l'ensemble des éléments $(u_h)_{h \in H}$ de \mathbb{N}^H tels que $\sum_{h \in H} u_h = n$.

On note que, à ce stade, il n'y a pas de différence "mathématique" entre les hypothèses données jusqu'à maintenant et celles données au paragraphe B qui précède. En effet, quand on se restreint aux macro-états (cf. la fin du 2°), **la différence essentielle entre le théorème de Jackson et le théorème BCMP se situe au niveau des taux de service**. La formalisation que nous proposons est un peu plus générale que celle du théorème BCMP classique : pour autant, nous verrons que la "démonstration" du théorème associé est quasi-triviale.

4°) Soit J une famille de parties de H ; pour tout élément h de H , soit $b(h)$ la famille des éléments de J dont h est un élément ; soit d une fonction positive définie sur $(J \times \mathbb{N})$ telle que $d(j, 0) = 0$ et, quel que soit $i > 0$, $d(j, i) > 0$.

Pour tout élément (j, h, u) de $(J \times H \times E)$ et pour tout entier k , on pose :

$$(4C1) \quad u''_j := \sum_{i \in j} u_i$$

$$(4C2) \quad s(h, u) := \prod_{j \in b(h)} d(j, u''_j)$$

$$(4C3) \quad d'(j,k) := \prod_{i \leq k} d(j,i)$$

5°) Si, **par exemple**, le taux de service des clients de classe i à la station j ne dépend que du nombre total de tels clients, et ceci pour tout couple (i,j) , J est l'ensemble des parties réduites à un point.

De même, si le taux de service de la classe i à la station j ne dépend que du nombre total de clients dans cette station j , et ceci pour tout couple (i,j) , à chaque station j correspond un et un seul élément de J qui est la partie $P(j)$ de H constituée de l'ensemble des couples (i,j) , j fixé, i quelconque.

Il y a aussi des cas où le taux de service de la classe i à la station j peut faire intervenir plusieurs nombres liés à l'état u : par exemple, supposons que "le" serveur ait un taux de service global égal à l'unité et qu'il se répartisse équitablement entre tous les clients de la station (temps partagé), c'est à dire que, si $h := (i,j)$, $s(h,u) = u_h / \sum_k u_{(k,i)} = u_{(i,j)} / \sum_k u_{(k,i)}$.

Pour simplifier l'expression, supposons qu'il en soit ainsi pour tous les couples (i,j) . Dans ce cas, J comprend à la fois les parties réduites à un point et, pour chaque station i , la partie $P(i)$ de H constituée des couples (j,i) , i fixé, j quelconque. De plus, dans ce cas, on pose :

pour chaque partie réduite à un point :

$$d(\{h\}, u_h'') = u_h \quad (\text{ici } u_h'' = u_h)$$

et pour chaque partie $P(i) := \bigcup_j \{j,i\}$:

$$d(P(i), u_{P(i)}'') = 1 / u_{P(i)}''$$

$$\text{où } u_{P(i)}'' := \sum_j u_{(j,i)}.$$

Ces trois exemples montrent que la formulation proposée au 4°) est très générale malgré son apparente simplicité. Il faut évidemment que l'ensemble E des états que l'on considère soit compatible avec les taux de service et les routages considérés, mais, là encore, il ne faut pas hésiter à prendre un ensemble d'états "surprenant".

C.2 Théorème BCMP ("macro-états")

Soit q' la fonction définie sur E par :

$$q'(u) := \left\{ \prod_{h \in H} x_h^{u(h)} \right\} / \left\{ \prod_{j \in J} d'(j, u_j'') \right\}$$

On pose $c := \sum_{u \in E} q'(u)$ et $q(u) := q'(u)/c$.

Alors q est la probabilité stationnaire.

Preuve :

Comme en B.2, on va prouver que, pour tout élément h de H , on a :

$$(4C4) \quad \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h})$$

Si H est l'ensemble des couples (i, j) où i est une classe de clients et j une station, et quand la relation 4C4 est satisfaite, on dit qu'il y a **balance locale par classe et par station**.

La preuve de 4C4 est très voisine de celle proposée en B.3. Plus précisément, on a :

$$\begin{aligned} \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h}) &:= \sum_{j \neq h} q(u + e_j - e_h) \prod_{i \in b(j)} d(i, u''_i + 1) r_{j,h} \\ &= \sum_{j \neq h} q(u) x_j \left\{ \prod_{i \in b(h)} d(i, u''_i) \right\} r_{j,h} / x_h \\ &= q(u) \left\{ \prod_{i \in b(h)} d(i, u''_i) \right\} =: \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) \end{aligned}$$

C.3 Remarques

1°) Comme on l'a déjà noté, le théorème démontré précédemment ne coïncide pas exactement avec le théorème BCMP classique. D'une part, la formulation concernant les taux de service proposée ici est plus abstraite et un peu plus générale. Notamment, à titre illustratif, on a donné comme exemple de base associé à l'hypothèse 4C2 le cas $h = (i, j)$ où i est une station et j une classe de clients. En fait, 4C2 généralise la construction des taux de service proposée en B.1 même s'il n'y a qu'une seule classe de clients ; qu'il y ait une ou plusieurs classes de clients, cette généralisation est presque indispensable quand diverses parties d'un système sont servies **par divers microprocesseurs dont chacun travaille en temps partagé** (cf. la fin de C.1). Par contre la formulation proposée en C.2 est insuffisante si les disciplines de service dans certaines stations prennent en compte l'ordre d'arrivées des clients (de classes différentes). Ce cas sera abordé en F.

2°) Le lecteur novice, notamment le lecteur mathématicien est peut-être surpris par le caractère élémentaire des démonstrations données dans ces paragraphes A, B et C : il a pourtant fallu plusieurs années pour "y penser".

3°) L'étude effectuée en C.2 s'étend sans difficulté au cas ouvert, comme en B.3. En fait, avec la formulation 4C1 (H est une partie de H), la différence entre le cas ouvert et le cas fermé est faible ; elle ne porte que sur E (cas E infini et cas E fini).

4°) Pour garder une formulation voisine de celle du théorème BCMP classique, on a considéré que se donner h correspond à une classe et une station. On pourrait supposer que $h = (h_1, h_2, h_3)$ où, par exemple, h_1 est le type du programme, h_2 le processeur qui le prend en charge et h_3 la tâche en cours d'exécution : si les "taux de service" admettent une forme produit, on a l'analogie du théorème 4.C.2. Ce point est un cas particulier du paragraphe D qui suit.

D. Modification des taux de service

D.1 Introduction

La suite de ce chapitre va consister à montrer comment on peut "construire" des réseaux de files d'attente dont on sait calculer la probabilité stationnaire à partir d'autres réseaux plus simples, notamment les réseaux qu'on a introduits aux trois paragraphes précédents (cf. [Len], [Pel-2] et [Pel-4]).

On a vu, au paragraphe C qui précède, que le passage du théorème de Jackson au théorème BCMP (quand on se restreint aux macro-états) porte sur une modification des taux de service : ce paragraphe va formaliser et généraliser ce point.

Plus précisément, le but de ce paragraphe D, relativement élémentaire, est double : d'une part, on y montre que, si q est la probabilité stationnaire d'un réseau (R, R') et s'il y a "balance locale" de R' par rapport à R (cf. D.2.8°), **on peut modifier les taux de service dans R'** , de façon multiplicative, sans perdre la propriété de balance locale, ni pour R' , ni éventuellement, pour une autre station A de R (ce qui permet de répéter le procédé) ; de plus, quand on effectue cette modification des taux de service, la nouvelle probabilité stationnaire q' s'exprime simplement en fonction de q .

D'autre part, ce paragraphe doit permettre au lecteur de se familiariser avec des notations telles que :

$$\tau(u, R), \tau(u, R), \text{ etc...}$$

Dans tout ce qui suit, le lecteur notera qu'on prend bien garde de distinguer d'une part les hypothèses mathématiques, d'autre part l'interprétation concrète qui en est donnée ; cette interprétation facilite la

compréhension pour le lecteur mais elle n'est jamais utilisée au niveau des démonstrations.

D.2 Hypothèses de base

1°) On considère un système markovien homogène (R, R') qui admet E comme ensemble d'états : E peut être fini ou infini. Soit a la "matrice" qui régit l'évolution de ce système comme rappelé en A.2.

2°) Soit n une fonction entière positive (au sens large) définie sur E ; soit s une fonction réelle positive définie sur \mathbb{N} avec $s(0) = 0$ et, pour $i > 0$, $s(i) > 0$. On suppose que $|n(u) - n(v)| > 1$ implique $a(u, v) = 0$.

3°) Donnons tout de suite **une interprétation** de ces fonctions n et s : n est le nombre de clients (en tout, pour une classe, pour plusieurs classes réunies, etc...) d'un certain type dans la partie R' du système considérée. La modification que l'on va étudier consiste à multiplier par $s[n(u)]$ les taux, pour les clients associés à n , de passage de R' dans R ; par exemple, R' peut être une station unique et on modifie les taux de service dans cette station pour certains clients.

4°) Soit q une fonction positive définie sur E qui satisfait aux propriétés suivantes :

$$\sum_{u \in E} q(u) < +\infty \quad \text{et} \quad \sum_{u \in E} \sum_{v \in E} q(u) a(u, v) < +\infty$$

et, pour tout élément u de E :

$$(4D1) \quad \sum_{v \in E} q(u) a(u, v) = \sum_{v \in E} q(v) a(v, u)$$

Autrement dit, à une constante multiplicative près, q peut être considérée comme une probabilité stationnaire pour le couple (E, a) .

5°) Pour tout élément u de E , on pose :

$$B(u) := \{ v : v \in E, n(v) = n(u) - 1 \}$$

$$C(u) := \{ v : v \in E, n(v) = n(u) \}$$

$$D(u) := \{ v : v \in E, n(v) = n(u) + 1 \}$$

$$\begin{array}{c} \uparrow \uparrow \\ \tau(u, R) = q(u) \sum_{v \in D(u)} a(u, v) \end{array}$$

$$\begin{array}{c} \downarrow \downarrow \\ \tau(u, R) = \sum_{v \in D(u)} q(v) a(v, u) \end{array}$$

6°) Les deux quantités qui précèdent peuvent être interprétées de la façon suivante (ce qui explique les notations utilisées) : la première est le "taux de probabilité" de "quitter" l'état u , un client (du type considéré) passant de R à R' (et donc quittant R) ; la deuxième est le "taux de probabilité" "d'atteindre" l'état u , un client (du type considéré) rentrant dans R (en venant de R').

7°) On définit de même :

$$\tau(u, R)^\uparrow := q(u) \sum_{v \in C(u)} a(u, v)$$

$$\tau(u, R)^\downarrow := \sum_{v \in C(u)} q(v) a(v, u)$$

$$\tau(u, R)^\uparrow \downarrow := q(u) \sum_{v \in B(u)} a(u, v)$$

$$\tau(u, R)^\downarrow \uparrow := \sum_{v \in B(u)} q(v) a(v, u)$$

$$\tau(u)^\uparrow := \tau(u, R)^\uparrow + \tau(u, R)^\uparrow \circ + \tau(u, R)^\uparrow \downarrow$$

$$\tau(u)^\downarrow := \tau(u, R)^\downarrow + \tau(u, R)^\downarrow \circ + \tau(u, R)^\downarrow \uparrow$$

On a évidemment (cf. D.2.2°) :

$$\tau(u)^\uparrow = q(u) \sum_{v \in E} a(u, v) \text{ et}$$

$$\tau(u)^\downarrow = \sum_{v \in E} q(v) a(v, u)$$

et la relation 4D1 s'écrit :

$$\tau(u)^\uparrow = \tau(u)^\downarrow$$

8°) On suppose que, pour tout élément u de E , on a :

$$(4D2) \quad \tau(u, R)^\uparrow \downarrow = \tau(u, R)^\downarrow \uparrow$$

Si on revient à l'interprétation proposée au 6°) qui précède et si R' est une station S unique, 4D2 est exactement la "balance locale" relativement à S (pour les clients du type considéré) ; en général, au niveau des applications, R' n'est pas une station unique.

D.3 Théorème de stabilité

On suppose que l'on a les hypothèses de base données au paragraphe D.2 qui précède, c'est à dire, essentiellement, les relations 4D1 et 4D2. Pour tout entier k , on pose :

$$t(k) := \prod_{j \leq k} s(j)$$

Pour tout élément u de E on pose :

$$q'(u) := q(u) / t[n(u)]$$

On suppose que :

$$\sum_{u \in E} q'(u) < +\infty$$

Enfin, pour tout couple (u, v) d'éléments de E , on pose :

- si $n(v) = n(u) - 1$, $a'(u, v) := a(u, v) s[n(u)]$
- sinon, $a'(u, v) := a(u, v)$

On définit $\tau'(\overset{\uparrow}{u})$, $\tau'(\overset{\uparrow}{u}, \overset{\downarrow}{R})$, etc... relativement au couple (q', a') exactement comme on a défini $\tau(\overset{\uparrow}{u})$, $\tau(\overset{\uparrow}{u}, \overset{\downarrow}{R})$, etc... relativement au couple (q, a) .

On suppose que $\tau'(\overset{\uparrow}{u}) < +\infty$.

On a alors, pour tout élément u de E :

$$(4D3) \quad \tau'(\overset{\uparrow}{u}) = \tau'(\overset{\downarrow}{u})$$

$$(4D4) \quad \tau'(\overset{\uparrow}{u}, \overset{\downarrow}{R}) = \tau'(\overset{\downarrow}{u}, \overset{\uparrow}{R})$$

Interprétation

Revenons à l'interprétation proposée en D.2.6°) ; si R' est une station S unique, a' se déduit de a en multipliant le taux de service dans S par $s[n(u)]$ où $n(u)$ est le nombre de clients (de la classe considérée) dans S . La relation 4D3 dit que q' est, à une constante multiplicative près, **la probabilité stationnaire associée à a'** . La relation 4D4 dit que l'on a encore balance locale pour S après cette modification.

Preuve :

1°) Soit u un élément fixé de E ; pour alléger les notations, on pose :

$$r := 1/t[n(u)-1]$$

Soit v un élément de $B(u)$ c'est à dire que $n(v) = n(u)-1$; on a alors :

$$a'(v,u) = a(v,u)$$

$$q'(v) = r q(v) \quad (\text{par définition de } q')$$

$$a'(u,v) = a(u,v) s[n(u)]$$

$$q'(u) = r q(u)/s[n(u)]$$

ce qui implique

$$q'(v) a'(v,u) = r q(v) a(v,u) \quad \text{et}$$

$$q'(u) a'(u,v) = r q(u) a(u,v)$$

ce qui implique

$$\begin{aligned} \tau'(\overset{\uparrow}{u}, \overset{\downarrow}{R}) &= q'(u) \sum_{v \in B(u)} a'(u,v) \\ &= r q(u) \sum_{v \in B(u)} a(u,v) = r \tau(\overset{\uparrow}{u}, \overset{\downarrow}{R}) \end{aligned}$$

De même :

$$\begin{aligned} \tau'(\overset{\downarrow}{u}, \overset{\uparrow}{R}) &= \sum_{v \in B(u)} q'(v) a'(v,u) \\ &= r \sum_{v \in B(u)} q(v) a(v,u) = r \tau(\overset{\downarrow}{u}, \overset{\uparrow}{R}) \end{aligned}$$

Les relations 4D2 impliquent donc les relations 4D4.

2°) L'état u est toujours fixé ; pour alléger les notations, posons :

$$w := r/s[n(u)] = 1/t[n(u)]$$

On a :

$$\tau'(\overset{\uparrow}{u}, \overset{\circ}{R}) = w \tau(\overset{\uparrow}{u}, \overset{\circ}{R})$$

$$\tau'(\overset{\downarrow}{u}, \overset{\circ}{R}) = w \tau(\overset{\downarrow}{u}, \overset{\circ}{R})$$

$$\tau'(\overset{\uparrow}{u}, \overset{\uparrow}{R}) = w \tau(\overset{\uparrow}{u}, \overset{\uparrow}{R})$$

En effet, pour tous les cas qui interviennent dans ces trois relations on a $q' = w q$ et $a' = a$.

Par ailleurs, si v appartient à $D(u)$, on a :

$$q'(v) = w q(v) / s[1+n(u)] \quad \text{et}$$

$$a'(v,u) = a(v,u) s[1+n(u)] \quad \text{d'où}$$

$$\tau'(\overset{\downarrow}{u}, \overset{\downarrow}{R}) = w \tau(\overset{\downarrow}{u}, \overset{\downarrow}{R})$$

En additionnant les relations qui précèdent, et en utilisant les définitions de $\tau'(\overset{\uparrow}{u})$, $\tau'(\overset{\downarrow}{u})$, $\tau(\overset{\uparrow}{u})$ et $\tau(\overset{\downarrow}{u})$ (cf. D.2.7°), on obtient :

$$\tau'(\overset{\uparrow}{u}) - \tau'(\overset{\uparrow}{u}, \overset{\downarrow}{R}) = w [\tau(\overset{\uparrow}{u}) - \tau(\overset{\uparrow}{u}, \overset{\downarrow}{R})] \quad \text{et}$$

$$\tau'(\overset{\downarrow}{u}) - \tau'(\overset{\downarrow}{u}, \overset{\uparrow}{R}) = w [\tau(\overset{\downarrow}{u}) - \tau(\overset{\downarrow}{u}, \overset{\uparrow}{R})]$$

Les deuxièmes membres de ces deux égalités sont égaux (relations 4D1 et 4D2) ; il en est donc de même pour les premiers membres, ce qui implique 4D3 compte tenu de 4D4 (cf. le 1° qui précède).

D.4 Conservation de la balance locale

On garde toutes les hypothèses introduites en D.2 et D.3. De plus, pour tout élément u de E , on considère deux parties $V(u)$ et $W(u)$ de E . On suppose que, pour tout élément u de E , on a :

- si v appartient à $V(u)$ ou $W(u)$, $n(v) \geq n(u)$

On suppose que l'on a :

$$(4D5) \quad q(u) \sum_{v \in V(u)} a(u,v) = \sum_{v \in W(u)} q(v) a(v,u)$$

Alors on a aussi :

$$(4D6) \quad q'(u) \sum_{v \in V(u)} a'(u,v) = \sum_{v \in W(u)} q'(v) a'(v,u)$$

Preuve :

Si v appartient à $V(u)$, $a'(u,v) = a(u,v)$, donc :

$$(4D7) \quad q'(u) \sum_{v \in V(u)} a'(u,v) = w q(u) \sum_{v \in V(u)} a(u,v)$$

Si v appartient à $W(u)$, il y a deux cas à considérer :

1°) v appartient à $W(u)$ et $C(u)$; dans ce cas, $q'(v) = w q(v)$ et $a'(v,u) = a(v,u)$;

2°) v appartient à $W(u)$ et $D(u)$; dans ce cas, on a :

$$a'(v,u) = a(v,u) / s [1+n(u)] \text{ et } q'(v) = w q(v) / s [1+n(u)]$$

Dans les deux cas (1°) et (2°), on a donc :
 $q'(v) a'(v,u) = w q(v) a(v,u)$.

Finalement, on arrive à :

$$\sum_{v \in W(u)} q'(v) a'(v,u) = w \sum_{v \in W(u)} q(v) a(v,u)$$

Cette relation et 4D7 reportées dans 4D5 donnent exactement la relation 4D6.

Interprétation

Le cas le plus simple est le cas où la relation 4D5 est la balance locale (pour une classe ou pour la réunion de plusieurs classes) relativement à une station A de R ; dans ce cas, les termes $\tau(u, \overset{\uparrow}{A})$ et $\tau(\overset{\downarrow}{u}, \overset{\downarrow}{A})$ ne font intervenir les coefficients $a(u,v)$ que pour des états v tels que $n(v) \geq n(u)$. Les hypothèses données en D.4 sont donc satisfaites.

Notamment, si, dans un réseau R, il y a balance locale pour les stations S et A, on peut modifier les taux de service dans S (comme indiqué précédemment) sans perdre ces propriétés de balance locale.

E. Produit de deux réseaux

E.1 Introduction

Ce paragraphe E est indispensable si on veut disposer d'un cadre suffisamment souple et général au niveau des applications.

L'idée essentielle de ce paragraphe est la suivante : on considère un réseau R' de files d'attente (en un sens très général) ; on suppose que ce réseau R' possède une station S' pour laquelle il y a balance locale par classe. Alors, on peut remplacer cette station S' par tout un réseau R'' sous réserve que le "noeud" de liaison entre R' et R'' puisse être considéré, relativement à R'', comme une station S'' pour laquelle il y a balance locale par classe (dans R'').

Quand on effectue cette substitution, la probabilité stationnaire du réseau composé (R',R'') est, à une constante multiplicative près, le produit des probabilités stationnaires associées à R' et R'' respectivement. De plus, s'il y a balance locale pour la station A, autre que S', dans R', il y a encore balance locale pour A dans le réseau composé (R',R'') : on peut donc, à nouveau, remplacer la station A par tout un réseau, etc... ; autrement dit,

on peut répéter ce procédé.

Par exemple, dans un réseau BCMP, soit R' , on peut remplacer une station par un réseau R'' à serveur central comme indiqué au paragraphe A ; puis, dans le réseau composé ainsi obtenu, on peut remplacer une station B de R'' par un réseau BCMP, etc... ; en général les réseaux que l'on obtient ainsi ne sont évidemment pas BCMP.

Du point de vue technique, les réseaux R' et R'' jouent des rôles parfaitement symétriques : **dans la suite, on va donc étudier le produit des réseaux R' et R''** ; l'hypothèse essentielle est que, dans le système (R', S') , où S' est une station unique, il y a balance locale pour S' et de même dans le système (R'', S'') . On supposera que, dans S' et S'' , tous les taux de service sont égaux à 1 ce qui allège notablement les notations et les démonstrations. Ceci n'est pas une restriction compte tenu du paragraphe D qui précède.

E.2 Hypothèses

Soit J un ensemble fini (cet ensemble J est le même pour R' et R''). Soit R' un réseau de files d'attente qui admet E' comme ensemble d'états ; soit a' la matrice qui régit l'évolution de R' . Pour tout élément j de J soit n'_j une fonction entière positive définie sur E' . On définit $B'(i, u)$, $C'(u)$, $D'(i, u)$, comme suit (cf. D.2) :

$$B'(i, u) := \{ v' : v' \in E', n'_i(v') = n'_i(u) - 1, \text{ et, pour } j \neq i, n'_j(v') = n'_j(u) \}$$

$$C'(u) := \{ v' : v' \in E', \text{ quel que soit } j, n'_j(v') = n'_j(u) \}$$

$$D'(i, u) := \{ v' : v' \in E', n'_i(v') = n'_i(u) + 1 \text{ et, pour } j \neq i, n'_j(v') = n'_j(u) \}$$

Soit q' une fonction positive définie sur E' . On définit $\tau'_i(\overset{\uparrow}{u'}, \overset{\uparrow}{R'})$, $\tau'_i(\overset{\downarrow}{u'}, \overset{\downarrow}{R'})$, etc... comme en D.2. Par exemple,

$$\tau'_i(\overset{\uparrow}{u'}, \overset{\uparrow}{R'}) := q'(u') \sum_{v' \in D'(i, u')} a'(u', v')$$

Soit s la fonction définie sur \mathbb{N} par $s(0) := 0$ et, pour $i > 0$, $s(i) := 1$. On suppose que l'on a, pour tout élément u' de E' :

$$(4E1) \quad \tau'_i(\overset{\uparrow}{u'}) = \tau'_i(\overset{\downarrow}{u'}) < +\infty$$

et, pour tout élément i de J :

$$(4E2) \quad q'(u') s(n'_i(u')) = \tau'_i(\overset{\downarrow}{u'}, \overset{\uparrow}{R'}) \quad \text{et}$$

$$(4E3) \quad s(n'_i(u')) = \sum_{v' \in B'(i, u')} a'(u', v')$$

Notons que l'hypothèse 4E2 est identique à 4D2 si S' est une "vraie" station dont le taux de service vaut 1 (condition 4E3) pour chaque classe de clients (s'il y a au moins un client de cette classe dans S').

Soit R'' un réseau de files d'attente, distinct de R' , qui satisfait au même type d'hypothèses que R' ; notamment, on appelle E'' son ensemble d'états et a'' la matrice qui définit son évolution; on se donne, pour tout élément j de J , la fonction n''_j , etc... On suppose que, pour tout élément (j, u'') de $(J \times E'')$ on a :

$$(4E4) \quad \tau''(\overset{\uparrow}{u''}) = \tau''(\overset{\downarrow}{u''}) \text{ et}$$

$$(4E5) \quad q''(u'') s(n''_i(u'')) = \tau''_i(\overset{\downarrow}{u''}, \overset{\uparrow}{R''})$$

$$(4E6) \quad s(n''_i(u'')) = \sum_{v'' \in B''(i, u'')} a''(u'', v'')$$

Soit n une fonction entière strictement positive définie sur J . Le produit R de R' et R'' est alors défini de la façon suivante : l'ensemble E des états est l'ensemble des couples $u := (u', u'')$, avec u' élément de E' et u'' élément de E'' , couples tels que, pour tout élément j de J ,

$$n'_j(u') + n''_j(u'') = n(j).$$

Pour tout couple (u, v) d'éléments de $(E' \times E'') \times (E' \times E'')$ avec $u := (u', u'')$ et $v := (v', v'')$, la matrice a qui régit l'évolution de R est définie comme suit :

si $u'' = v''$ et si, pour tout élément j de J , $n'_j(u') = n'_j(v')$, on pose $a(u, v) := a'(u', v')$;

si $u' = v'$ et si, pour tout élément j de J , $n''_j(u'') = n''_j(v'')$, on pose $a(u, v) := a''(u'', v'')$;

si il existe j élément de J tel que, pour $k \neq j$, $n_k(u') = n_k(v')$,

$n_k(u'') = n_k(v'')$ et, soit $\{ n'_j(v') = n'_j(u') + 1 \text{ et } n''_j(v'') = n''_j(u'') - 1 \}$, soit

$\{ n'_j(v') = n'_j(u') - 1 \text{ et } n''_j(v'') = n''_j(u'') + 1 \}$, alors on pose

$$a(u, v) := a'(u', v') a''(u'', v'') .$$

Dans tous les autres cas, on pose $a(u, v) := 0$.

Interprétation

On suppose que $n_j(u')$ est le nombre de clients de classe j dans S' quand l'état est u' ; $a(u,v)$ est différent de 0 dans chacun des quatre cas suivants (et seulement ceux-là) :

Premier cas : l'état de R' ne change pas, seul celui de R'' change, mais il n'y a ni entrée ni sortie de clients de R'' .

Deuxième cas : identique au précédent sauf qu'on intervertit R' et R'' .

Troisième cas : un client de classe j "quitte" la nébuleuse R' et "rentre" dans la nébuleuse R'' .

Quatrième cas : identique au précédent en intervertissant R' et R'' .

Pour le premier cas (resp. deuxième), le "taux d'évolution" dans R est le même que dans R' (resp. R''). Pour le troisième et le quatrième cas, le "taux d'évolution" dans R est le produit du taux d'évolution dans R' et du taux d'évolution dans R'' .

On peut aussi présenter l'évolution de la façon suivante : le réseau global R est constitué de R' , R'' , S' et S'' , les clients ne restant pas dans S' et S'' (taux de service "infinis" en S' et S''). Plus précisément, un client qui arrive dans S' (resp. S'') avec la classe j va aussitôt dans S'' (resp. S') sans changer de classe et est immédiatement servi dans S'' (resp. S'), et rentre donc dans R'' (resp. R') avec les "routages" associés au couple (R'', S'') (resp. (R', S')).

Attention : le départ d'un client de classe j de R' peut être dû au départ d'un client de classe k de la station A de R' , client qui se transforme en client de classe j avant de quitter R' ; de même, un client de classe j qui rentre dans R'' peut aussitôt être transformé en client de classe k et aller dans une station B de R'' , ou même provoquer plusieurs modifications conjointes.

E.3 Théorème de composition

On considère les hypothèses données au paragraphe E.2 qui précède. Pour tout élément $u = (u', u'')$ de E , on pose $q(u) = q'(u') q''(u'')$. On suppose que :

$$\sum_{u \in E} q(u) < +\infty \quad \text{et} \quad \tau(\overset{\uparrow}{u}) < +\infty .$$

On a alors :

$$(4E7) \quad \tau(\overset{\uparrow}{u}) = \tau(\overset{\downarrow}{u})$$

où, évidemment, ces notations sont définies comme précédemment c'est à

dire que :

$$\begin{aligned}\tau(\overset{\uparrow}{u}) &:= q(u) \sum_{v \in E} a(u,v) \quad \text{et} \\ \tau(\overset{\downarrow}{u}) &:= \sum_{v \in E} q(v) a(v,u)\end{aligned}$$

Autrement dit q peut être considérée, à une constante multiplicative près, comme la probabilité stationnaire associée au couple (E,a) .

Preuve :

Toutes les quantités considérées sont positives : il suffit donc de prouver $\tau(\overset{\downarrow}{u}) = \tau(\overset{\uparrow}{u})$ quand l'ensemble des états est fini.

On a $\tau(\overset{\downarrow}{u}) = z' + z''$ avec

$$\begin{aligned}z' &:= q'(u') \tau''(\overset{\downarrow}{u''}, \overset{\circ}{R''}) + \sum_{j \in J} \tau_j'(\overset{\downarrow}{u'}, \overset{\uparrow}{R'}) \tau_j''(\overset{\downarrow}{u''}, \overset{\downarrow}{R''}) \\ z'' &:= q''(u'') \tau'(\overset{\downarrow}{u'}, \overset{\circ}{R'}) + \sum_{j \in J} \tau_j''(\overset{\downarrow}{u''}, \overset{\uparrow}{R'}) \tau_j'(\overset{\downarrow}{u'}, \overset{\downarrow}{R'})\end{aligned}$$

Compte tenu de 4E2, on a ($n_j'(u') = 0$ implique $\tau_j''(\overset{\downarrow}{u''}, \overset{\downarrow}{R''}) = 0$) :

$$\begin{aligned}z' &= q'(u') \{ \tau''(\overset{\downarrow}{u''}, \overset{\circ}{R''}) + \sum_{j \in J} \tau_j''(\overset{\downarrow}{u''}, \overset{\downarrow}{R''}) \} \\ &= q'(u') \{ \tau''(\overset{\downarrow}{u''}) - \sum_{j \in J} \tau_j''(\overset{\downarrow}{u''}, \overset{\uparrow}{R''}) \}\end{aligned}$$

Compte tenu de 4E5 ceci implique :

$$= q'(u') \{ \tau''(\overset{\downarrow}{u''}) - q''(u'') \sum_{j \in J} s(n_j''(u'')) \}$$

On prouve de même que

$$z'' = q''(u'') \{ \tau'(\overset{\downarrow}{u'}) - q'(u') \sum_{j \in J} s(n_j'(u')) \}$$

Compte tenu de 4E1 et 4E4 on a donc

$$\begin{aligned}\tau(\overset{\downarrow}{u}) &= z' + z'' = q'(u') \tau''(\overset{\uparrow}{u''}) + q''(u'') \tau'(\overset{\uparrow}{u'}) \\ &\quad - q'(u') q''(u'') \sum_{j \in J} [s(n_j'(u')) + s(n_j''(u''))] \\ &= \tau(\overset{\uparrow}{u}) \quad (\text{par définition de } \tau(\overset{\uparrow}{u}) : \text{cf. 4E3 et 4E6}).\end{aligned}$$

On a donc prouvé 4E7 .

E.4 Conservation de la balance locale

On garde toutes les hypothèses introduites en E.2 et E.3. De plus, pour tout élément u' de E' , on considère deux parties $V'(u')$ et $W'(u')$ de E' . On suppose que, pour tout élément (j, u') de $(J \times E')$, on a :

si v' est un élément de $V'(u')$ ou de $W'(u')$, $n'_j(v') \geq n'_j(u')$

On suppose que l'on a :

$$(4E8) \quad q'(u') \sum_{v' \in V'(u')} a'(u', v') = \sum_{v' \in W'(u')} q'(v') a'(v', u')$$

Pour tout élément $u = (u', u'')$ de E , on définit $V(u)$ et $W(u)$ comme suit :

$V(u)$ est l'ensemble des couples $v = (v', v'')$ tels que v' appartient à $V'(u')$ et $a(u, v) \neq 0$; de même, $W(u)$ est l'ensemble des couples $w = (w', w'')$ tels que w' appartient à $W'(u')$ et $a(w, u) \neq 0$.

On a alors :

$$(4E9) \quad q(u) \sum_{v \in V(u)} a(u, v) = \sum_{v \in W(u)} q(v) a(v, u)$$

Preuve :

Nous allons supposer – ce qui n'est pas une restriction – que si v' appartient à $V'(u')$ (resp. $W'(u')$), $a'(u', v')$ (resp. $a'(v', u')$) est différent de 0.

On suppose que 0 n'appartient pas à J (ceci n'est qu'une commodité de notation) et on pose $J' := J \cup \{0\}$. On pose :

$$V'(u', 0) := \{v' : v' \in V'(u') \text{ et, } \forall k, n'_k(v') = n'_k(u')\}$$

$$V'(u', k) := \{v' : v' \in V'(u') \text{ et } n'_k(v') = 1 + n'_k(u')\}$$

De même, on pose :

$$W'(u', 0) := \{v' : v' \in W'(u') \text{ et, } \forall k, n'_k(v') = n'_k(u')\}$$

$$W'(u', k) := \{v' : v' \in W'(u') \text{ et } n'_k(v') = 1 + n'_k(u')\}$$

La famille $(V'(u', k))_{k \in J'}$ (resp. $(W'(u', k))_{k \in J'}$) constitue une partition de $V'(u')$ (resp. $W'(u')$). Notons aussi que $V'(u', k)$ ou $W'(u', k)$ non vide implique $n'_k(u') < n(k)$ et donc $n''_k(u'') > 0$ si (u', u'') appartient à E .

Soit $V(u,k)$ (resp. $W(u,k)$) l'ensemble des éléments $v = (v',v'')$ de $V(u)$ (resp. $W(u)$) tels que v' appartient à $V'(u,k)$ (resp. $W'(u,k)$).

Appelons x (resp. y) le premier (resp. deuxième) membre de 4E9. On a :

$$x = q(u) \sum_{k \in J'} \sum_{v \in V(u,k)} a(u,v) \text{ et}$$

$$y = \sum_{k \in J'} \sum_{v \in W(u,k)} q(v) a(v,u)$$

Or, pour $k \neq 0$, v élément de $V(u,k)$ et $n_k''(v'') = n_k''(u'') - 1$, on a $a(u,v) = a'(u',v') a''(u'',v'')$ donc

$$\sum_{v \in V(u,k)} a(u,v) = \left\{ \sum_{v' \in V'(u',k)} a'(u',v') \right\} \left\{ \sum_{v'' \in B''(k,u'')} a''(u'',v'') \right\}$$

où $B''(k,u'')$ est défini comme en E.2, c'est à dire que

$$B''(k,u'') := \{ v'' : v'' \in E'', n_k''(v'') = n_k''(u'') - 1 \}$$

Autrement dit (cf. 4E6) :

$$q(u) \sum_{v \in V(u,k)} a(u,v) = \left\{ q'(u') \sum_{v' \in V'(u',k)} a'(u',v') \right\} \tau_k''(\overset{\uparrow}{u''}, \overset{\downarrow}{R''})$$

$$= \left\{ q'(u') \sum_{v' \in V'(u',k)} a'(u',v') \right\} q''(u'') s[n_k''(u'')]$$

avec $s[n_k''(u'')] = 1$ si $V'(u',k)$ n'est pas vide.

On a donc :

$$x = \sum_{k \in J'} \left\{ q'(u') \sum_{v' \in V'(u',k)} a'(u',v') \right\} q''(u'')$$

$$(4E10) \quad x = \left\{ q'(u') \sum_{v' \in V'(u')} a'(u',v') \right\} q''(u'')$$

Par ailleurs, pour $k \neq 0$ et $W'(u',k)$ non vide, on a :

$$\sum_{v \in W(u,k)} q(v) a(v,u) = \left\{ \sum_{v' \in W'(u',k)} q'(v') a'(v',u') \right\} \tau_k''(\overset{\downarrow}{u''}, \overset{\uparrow}{R''})$$

$$= \left\{ \sum_{v' \in W'(u',k)} q'(v') a'(v',u') \right\} q''(u'') s[n_k''(u'')]$$

(compte tenu de la relation 4E5) avec $s[n_k''(u'')] = 1$

On a donc :

$$y = \left\{ \sum_{k \in J'} \sum_{v' \in W'(u', k)} q'(v') a'(v', u') \right\} q''(u'')$$

$$y = \left\{ \sum_{v' \in W'(u')} q'(v') a'(v', u') \right\} q''(u'')$$

Les relations 4E8 et 4E10 impliquent alors $x = y$, c. q. f. d.

F. Exemples

F.1 Un seul client et corollaires

1°) Dans ce cas, on prend $E=H$ où H est, par exemple, l'ensemble des stations : un état u est caractérisé par la station dans laquelle se trouve le seul client du réseau. Relativement à la probabilité stationnaire on a évidemment, pour tout élément (u, h) de $(E \times H)$, $\tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h})$ puisque, ou bien ces quantités sont nulles (si $u \neq h$), ou bien (si $u=h$)

$$(4F1) \quad \tau(\overset{\uparrow}{u}, \overset{\uparrow}{h}) = \tau(\overset{\uparrow}{u}) = \tau(\overset{\downarrow}{u}) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{h})$$

On a donc balance locale pour chaque station.

2°) Soit s un élément de H et posons $H' := H \setminus \{s\}$. Considérons l'interprétation suivante : s est une station de service "externe" et H' est l'ensemble des états fictifs associé à une loi non exponentielle (cf. 1.F). **Puisqu'il y a balance locale en s , on peut composer H' et un autre réseau pour lequel il y a aussi balance locale (cf. le théorème E.3).**

3°) Considérons le serveur central introduit au paragraphe A.7 dans le cas particulier où l'ensemble E est l'ensemble $\{0,1\}^H$ (autrement dit, pour chaque station h , $u_h = 0$ ou $u_h = 1$ et $r(h,u) = 0$ si $u=h$) : puisque, pour chaque station h , il y a au plus un client et il y a balance locale pour cette station, on peut remplacer, dans le réseau à serveur central, une station par le réseau H' introduit au 2°) qui précède. Ceci ne modifie pas la propriété de balance locale pour les autres stations (cf. E.4) : on peut donc faire de même pour chaque station. **Autrement dit, dans le serveur central, on peut supposer que la loi de service dans chaque station est une loi quelconque** (que l'on modélise à l'aide d'états fictifs pour rendre le système markovien).

Evidemment, au niveau du serveur central, la décomposition en

stations dont la capacité est limitée à un client est en général une décomposition fictive ; le point important est que les hypothèses soient compatibles avec cette décomposition fictive. Notamment, il faut qu'il y ait suffisamment de stations pour accueillir tous les clients qui arrivent ; par contre, les taux de service dans chacune de ces stations ne contenant qu'une place peuvent, par exemple, dépendre multiplicativement du nombre total de clients d'une certaine classe dans l'ensemble du système (cas à "temps partagé").

4°) Considérons un réseau BCMP, ou, plus précisément, un réseau comme défini au paragraphe C. Dans un tel réseau, pour chaque station h (ou chaque couple (h, h'') , etc...) il y a balance locale ; on peut donc (théorème E.3) remplacer cette station par l'ensemble des stations internes du serveur central ; ensuite, on procède comme au 3°) qui précède.

Autrement dit, dans un réseau comme défini au paragraphe C, on peut supposer que pour certaines stations la loi de service est non exponentielle (dans ce cas, chaque client commence à être servi dès qu'il est entré dans une telle station) : ceci préserve la balance locale pour les autres stations et la forme produit de la solution stationnaire. L'ensemble des états prend en compte les états fictifs associés à ces lois de service : on dit qu'on considère les micro-états.

F.2 Premier arrivé, premier servi

1°) Considérons un système R' - avec plusieurs classes de clients - dont on "sait" calculer la probabilité stationnaire q' . Soit S' une station (comme définie aux paragraphes précédents) pour laquelle il y a balance locale (les lois de service étant exponentielles). Alors, on peut remplacer S' par une station R pour laquelle la discipline de service est "premier arrivé, premier servi", (on dit aussi Fifo : first in, first out), sous réserves que les taux de service soient les mêmes pour chaque classe de clients : la probabilité stationnaire associée au réseau (R', R) est (à une constante multiplicative près) le produit de la probabilité q' et de la probabilité q associée à R .

On va prouver ce résultat en utilisant le théorème E.3. Il suffit donc de montrer que, dans le système (R, S) où S est une station dont le taux de service égale 1 pour chaque classe de clients, la propriété de balance locale est satisfaite pour S .

2°) On définit alors l'ensemble E des états (les "micro-états") de R de la façon suivante : K est un ensemble fini (c'est l'ensemble des classe de clients) ; un élément $u := (u_0, u_1, \dots)$ est une séquence où $u_0 = u(0)$ est un entier positif (c'est le nombre total de clients dans R) et $(u_i)_{1 \leq i \leq u(0)}$ est une séquence d'éléments de K , u_i étant la classe du i -ième client arrivé dans R (sans tenir compte des clients qui sont déjà repartis).

On suppose que le taux de service dans R vaut 1 quel que soit la classe du client en tête : seul ce client peut être servi. Un client servi dans R va dans S sans changer de classe. Dans S et pour chaque classe de clients le taux de service vaut 1 si S contient au moins un client de cette classe. Le réseau est fermé. Soit n_k le nombre total de clients de la classe k dans le réseau (n_k est fixe) : u caractérise l'état de (R,S). Attention : ce réseau (R,S) n'est qu'un intermédiaire technique, on pourra ensuite modifier les taux de service comme indiqué au paragraphe D.

La probabilité stationnaire est alors la probabilité équidistribuée ; soit q cette probabilité. On a :

quel que soit (u,k) élément de $(E \times K)$

$$(4F2) \quad \tau_k(\overset{\uparrow}{u}, \overset{\uparrow}{S}) = q(u) = \tau_k(\overset{\downarrow}{u}, \overset{\downarrow}{S})$$

(où $\tau_k(\overset{\uparrow}{u}, \overset{\uparrow}{S})$ est le "taux de probabilité" de quitter l'état par départ de S d'un client de classe k) et

$$(4F3) \quad \tau(\overset{\uparrow}{u}, \overset{\uparrow}{R}) = q(u) = \tau(\overset{\downarrow}{u}, \overset{\downarrow}{R})$$

Ceci montre que q est bien la probabilité stationnaire et qu'il y a balance locale par classe pour la station S. On peut donc effectuer le "produit" des réseaux R' et R (comme indiqué en E.3).

3°) Comme on l'a déjà noté, les taux de service peuvent être modifiés "globalement". Par contre, la restriction importante dans ce qui précède est que le taux de service doit être le même pour toutes les classes de clients.

4°) **En conclusion, dans un réseau défini à partir des macro-états, on peut supposer que, pour certaines stations – satisfaisant initialement à la propriété de balance locale et dont les taux de service ne dépendent pas de la classe des clients considérés – la discipline de service est premier arrivé, premier servi : ceci n'altère pas la propriété de balance locale pour d'autres stations et la forme produit de la probabilité stationnaire.**

Chapitre 5

Matrices

A Analyse numérique

A.1 Notations matricielles

1°) Les notations matricielles de base ont déjà été utilisées à plusieurs reprises dans ce cours (cf. notamment, 2.A.1). Rappelons seulement que, si m et n sont deux entiers strictement positifs, une **matrice réelle** $(m \times n)$ est une famille $a := (a_{i,j})$ de réels avec $1 \leq i \leq m$ et $1 \leq j \leq n$. On dit que $a_{i,j}$ est le **terme** de la i -ième ligne et de la j -ième colonne. Si a et b sont deux matrices $(m \times n)$ et $(n \times p)$ respectivement, $a \cdot b$ est la matrice $(m \times p)$ dont le terme général est défini par

$$(a \cdot b)_{i,j} := \sum_{k=1}^n a_{i,k} b_{k,j}$$

2°) Soit a une matrice $(m \times n)$; soit m' et n' deux entiers avec $1 \leq m' \leq m$ et $1 \leq n' \leq n$; soit b (resp. c, d, e) la matrice dont le terme général est défini comme suit

$$\begin{aligned} b_{i,j} &:= a_{i,j} && \text{pour } 1 \leq i \leq m' \text{ et } 1 \leq j \leq n' \\ c_{i,j} &:= a_{i,n'+j} && \text{pour } 1 \leq i \leq m' \text{ et } 1 \leq j \leq n-n' \\ d_{i,j} &:= a_{m'+i,j} && \text{pour } 1 \leq i \leq m-m' \text{ et } 1 \leq j \leq n' \\ e_{i,j} &:= a_{m'+i,n'+j} && \text{pour } 1 \leq i \leq m-m' \text{ et } 1 \leq j \leq n-n' \end{aligned}$$

On note alors symboliquement la matrice a de la façon suivante :

$$a = \begin{pmatrix} b & c \\ d & e \end{pmatrix}$$

et on dit que l'on a **décomposé a par blocs**. Evidemment, on peut être amené à considérer d'autres décompositions par blocs plus complexes que celle qui précède.

3°) Aux 1°) et 2°) ci-dessus, on a supposé que les indices i et j parcourent les ensembles totalement ordonnés $\{1, \dots, m\}$ et $\{1, \dots, n\}$; en fait, en général, pour les études effectuées dans ce cours, **les relations d'ordre sur les ensembles d'indices n'interviennent pas** en tant que relations

d'ordre total mais en tant que relations d'ordre partiel.

Par exemple, au niveau de la présentation, l'écriture

$$a = \begin{pmatrix} b & c \\ d & e \end{pmatrix}$$

utilise la relation d'ordre total mais, sur le fond, la décomposition par blocs ne fait pas intervenir de relation d'ordre sur les ensembles d'indices.

Plus précisément, étant donnés deux ensembles finis E et F , on dira que a est une **"matrice réelle généralisée" indexée par $(E \times F)$** si $a := (a_{u,v})$ est une famille de réels avec u élément de E et v élément de F . Si a (resp. b) est une matrice réelle indexée par $(E \times F)$ (resp. $(F \times G)$), la matrice a b est indexée par $(E \times G)$ et son terme général est défini (pour (u,w) élément de $(E \times G)$) par :

$$(a \ b)_{u,w} := \sum_{v \in F} a_{u,v} b_{v,w}$$

Si (E', E'') (resp. (F', F'')) est une partition de E (resp. de F), E' , E'' , F' et F'' n'étant pas vides, on peut décomposer a par blocs en posant :

$$b_{u,v} := a_{u,v} \quad \text{pour } (u,v) \text{ élément de } (E' \times F')$$

$$c_{u,v} := a_{u,v} \quad \text{pour } (u,v) \text{ élément de } (E' \times F'')$$

$$d_{u,v} := a_{u,v} \quad \text{pour } (u,v) \text{ élément de } (E'' \times F')$$

$$e_{u,v} := a_{u,v} \quad \text{pour } (u,v) \text{ élément de } (E'' \times F'')$$

On dit que a est une **matrice carrée** si $E = F$.

On dit que a est une **matrice diagonale** si a est une matrice carrée telle que $u \neq v$ implique $a_{u,v} = 0$.

4°) Soit w une matrice carrée. On dit que w est **monotone** si w est **inversible** et si son inverse w^{-1} est une **matrice positive**. On dit que w est une **L-matrice** si $w_{i,i} \geq 0$ (quel que soit i) et, quels que soient i et j avec $i \neq j$, $w_{i,j} \leq 0$. Par exemple, dans la relation 5A3 ci-dessous, la matrice $(d-a)$ est une L-matrice. Une **L-matrice monotone** est une **M-matrice**. Pour certains auteurs une M-matrice n'est pas nécessairement inversible : nous utiliserons donc systématiquement l'expression "M-matrice inversible".

5°) Soit a une matrice carrée. On dit que a est **acyclique** s'il n'existe pas de séquence $(u_k)_{1 \leq k \leq n}$ d'éléments de E telle que $u_1 = u_n \neq u_2$ et, quel que

soit k , $0 \leq k < n$, $a(u_k, u_{k+1}) \neq 0$. Ceci équivaut à dire que a est **triangulaire** pour un choix adéquat de la relation d'ordre sur l'ensemble E des indices.

Dans tout ce qui suit, on fera apparaître des matrices dont certaines possèdent les propriétés que l'on vient d'indiquer.

A.2 Hypothèses générales

On se donne un ensemble fini E et une fonction positive a définie sur $(E \times E)$; suivant les conventions habituelles dans ce cours, on pose

$a_{u,v} := a(u,v)$; **a est une matrice carrée positive** au sens généralisé donné en A.1.3°) indexée par $(E \times E)$. On suppose que, pour tout élément u de E , $a_{u,u} = 0$.

L'essentiel de ce cours (cf. 2.D.1) consiste à étudier et à déterminer les familles de nombres positifs $(q_e)_{e \in E}$ (les "matrices" unilignes) satisfaisant aux deux conditions suivantes :

(5A1) quel que soit e élément de E

$$q_e \sum_{v \in E} a_{e,v} = \sum_{v \in E} q_v a_{v,e}$$

(5A2) $\sum_{e \in E} q_e = 1$

Soit d la matrice "diagonale" $(E \times E)$ définie par $d_{u,v} := 0$ si $u \neq v$ et, pour $u = v$:

$$d_{u,u} := \sum_{v \in E} a_{u,v}$$

La famille de relations 5A1 équivaut à la relation matricielle suivante :

(5A3) $q d = q a$

La liaison entre 5A3 et l'étude des processus a été résumée en 2.D.2.

En fait, on s'intéresse essentiellement au cas où la matrice a est **irréductible** (cf. 2.B.3) ; dans ce cas (cf. 2.D.3), on sait qu'il existe une seule matrice q telle que l'on ait à la fois 5A2 et 5A3 ; chaque terme de cette matrice q est **strictement positif**. Toutes les remarques effectuées en 2.D.4 montrent que l'étude des solutions du système 5A3 relève, d'un certain point de vue, de l'analyse numérique.

Plus précisément, l'étude du système 5A3 peut être considérée comme un chapitre de la résolution des systèmes d'équations linéaires. Notamment, chacune des méthodes générales de résolution de tels systèmes peut être la plus efficace dans un contexte précis : **ces méthodes de base doivent donc servir de référence** (cf. [Cia], [GoM], [GoV], [PSS], [Ste], [Var], [Ch2], etc...). Toutefois, le système 5A3 a diverses propriétés spécifiques: cette spécificité est d'ailleurs renforcée si on considère des cas particuliers associés à des situations concrètes.

Enfin, assez curieusement, la résolution du système 5A3 fait parfois apparaître des situations analogues à celles que l'on a l'habitude de rencontrer en analyse numérique classique mais avec des variantes.... inhabituelles !

A.3 Remarques générales

1°) Même si certaines des méthodes proposées dans la suite de ce chapitre sont inhabituelles en analyse numérique classique, nous verrons qu'elles utilisent toutes très fortement le caractère **linéaire** du système 5A3.

2°) En général, les matrices a qui sont associées à des exemples concrets sont très fortement **creuses** (un très grand nombre des termes a_{ij} sont nuls). Il est donc impératif d'utiliser des techniques adaptées à cette situation (cf. par exemple, [DER], [Pis], etc...). De plus le **graphe** associé aux termes non nuls a une structure **très précise** mais aussi **très complexe** et qui dépend profondément des exemples traités.

3°) On a déjà noté (cf. 2.D.3 et 2.D.4) que résoudre le système 5A3 équivaut à résoudre le système $x = x r$ avec $x = q d$ et $r = d^{-1} a$; on a donc à déterminer le vecteur propre associé à la valeur propre de module maximum. Pour autant, les méthodes associées à l'étude théorique (cf. 2.B : calcul de r^k ou calculs analogues) sont tout à fait inadaptées ; on sait qu'elles convergent mais beaucoup trop lentement ; de plus, elles font intervenir des matrices "pleines".

Par contre, nous verrons plus loin qu'il est souvent possible de faire apparaître des produits de matrices positives (§ E.3) et chapitre 7), comme en analyse numérique classique.

4°) On a vu en 1.F qu'il fallait souvent introduire des "états fictifs" pour obtenir une modélisation markovienne. Au niveau des applications, en général, seules les probabilités "marginales" associées aux états "réels" sont intéressantes ; il suffit même souvent de ne garder en mémoire que **quelques combinaisons linéaires de ces "marginales"**. Il faut donc - pour diminuer la place mémoire utilisée - privilégier les algorithmes qui calculent les termes de q "par morceaux" et non globalement (cf. B.5,3°) et [Rob]).

5°) Evidemment, presque toutes les remarques générales liées à la résolution des grands systèmes linéaires restent valables dans le cas particulier du système 5A3. Notamment, il est souvent opportun de privilégier les algorithmes adaptés au calcul parallèle (cf. C.2,3°).

De plus, la stabilité de la méthode de résolution est un (le) problème crucial. Rappelons que **l'instabilité est liée en partie au nombre total d'opérations à effectuer mais surtout au rôle des "soustractions"** : l'erreur relative sur $(u-v)$ peut alors être beaucoup plus grande que la somme des erreurs relatives sur u et v respectivement, contrairement aux autres opérations sur des nombres positifs u et v (cf. B.2,2°, B.4,2°, le paragraphe D et E.5 in fine).

6°) Il ne faut pas oublier ce qui a été noté lors des chapitres 3 et 4 qui précèdent. D'une part, il serait absurde d'utiliser une méthode numérique matricielle générale si les hypothèses données au chapitre 4 sont "presque" satisfaites. D'autre part (cf. chapitre 3), il est presque toujours **"facile"** d'écrire le **programme de simulation** associé au système 5A3 : une étude numérique n'a donc d'intérêt que si elle est - ou peut devenir - plus performante (au moins d'un certain point de vue) que la simulation associée.

7°) **La matrice a** (et donc la matrice $(a-d)$) **est presque toujours tridiagonale par blocs**, et ceci pour plusieurs indexations distinctes : la raison en sera donnée en C.1. On va donc étudier tout particulièrement cette situation en commençant (§ B) par le cas où a est une matrice tridiagonale : évidemment, le but essentiel de ce paragraphe B est de préparer les paragraphes suivants et il doit être lu dans cette perspective.

8°) Pour alléger la présentation et utiliser l'écriture matricielle, dans ce qui suit on suppose en général que l'ensemble des états est fini. Evidemment, tout ce qui suit s'étend sans difficulté au cas où l'ensemble des états est infini quand les familles qui interviennent sont sommables. Cette extension sera formalisée à titre illustratif dans le paragraphe E.6.

B Matrices tridiagonales

B.1 Cadre général

1°) Soit m un entier, $m > 1$. On va étudier le **système linéaire $y t = x$** où t est une matrice $(m \times m)$ et x et y sont deux matrices unilignes $(1 \times m)$: **y est la matrice des "inconnues" tandis que x est supposée "connue"**. Dans tout ce paragraphe B, on suppose que t est **tridiagonale** c'est à dire que $|i-j| > 1$ implique $t_{i,j} = 0$. Pour alléger les notations, on pose :

$$\text{pour } 1 \leq i \leq m, \quad v_i := t_{i,i}, \quad x_i := x_{1,i} \quad \text{et} \quad y_i := y_{1,i}$$

pour $2 \leq i \leq m$, $u_i := -t_{i-1,i}$

pour $1 \leq i \leq m-1$, $w_i := -t_{i+1,i}$

2°) Le système $y \, t = x$ équivaut alors au système suivant d'équations :

$$y_1 v_1 - y_2 w_1 = x_1, \quad -y_{m-1} u_m + y_m v_m = x_m \text{ et,}$$

$$\text{pour } 1 < i < m, \quad -y_{i-1} u_i + y_i v_i - y_{i+1} w_i = x_i$$

3°) Si $y = q$, $x = 0$ et $t = (d-a)$, le système $y \, t = x$ devient $q \, d = q \, a$ (cf. 5A3) ; dans ce cas (cf. B.3 ci-dessous) $v_i = w_{i-1} + u_{i+1}$ (par définition de d) et tous les coefficients u_i , v_i et w_i sont positifs.

Toutefois, nous verrons plus tard que, pour résoudre le système général $q \, d = q \, a$, on peut être amené, comme étape intermédiaire, à résoudre des systèmes de la forme $y \, t = x$ avec $x \neq 0$ et $v_i \neq w_{i-1} + u_{i+1}$. En général on a $v_i \geq w_{i-1} + u_{i+1}$. De plus, même si on se restreint à considérer le système $y \, t = 0$, il est intéressant d'étudier la stabilité du système $y \, t = x$. On notera le rôle essentiel de la relation d'ordre sur l'ensemble des indices.

B.2 La factorisation $t = t' t''$

1°) Une méthode classique (cf., par exemple, [Cia]) de résolution du système $y \, t = x$ consiste à déterminer les matrices bidiagonales (et donc faciles à inverser) t' et t'' telles que $t = t' t''$, $t'_{i,i} = 1$, $t'_{i,j} = 0$ pour $j > i$ ou $j+1 < i$, et $t''_{i,j} = 0$ pour $j < i$ ou $j > i+1$. Evidemment, on ne met pas en mémoire t' et t'' sous forme matricielle.

Plus précisément, on construit successivement les trois suites suivantes (récurrence croissante pour les deux premières suites, récurrence décroissante pour la troisième suite) :

$$z_1 = 1/v_1 ; \quad z_j = 1/(v_j - u_j w_{j-1} z_{j-1}) \quad \text{pour } 2 \leq j \leq m$$

$$r_1 = z_1 x_1 ; \quad r_j = z_j (x_j + u_j r_{j-1}) \quad \text{pour } 2 \leq j \leq m$$

$$y_m = r_m ; \quad y_j = r_j + z_j w_j y_{j+1} \quad \text{pour } m > j \geq 1$$

La famille y est la solution cherchée.

2°) Supposons que, quel que soit j , les coefficients x_j , u_j , v_j et w_j soient positifs et que $v_j > u_{j+1} + w_{j-1}$.

On vérifie facilement en raisonnant par récurrence croissante sur j que l'on a $0 \leq z_j \leq 1/u_{j+1}$. Dans tous les calculs qui précèdent, il n'y a donc de "soustractions" qu'au niveau de la détermination des coefficients positifs $(v_j - u_j w_{j-1} z_{j-1})$ - coefficients qui ne font pas intervenir les variables x et y (cf. la remarque A.3, 5°). Pour tous les autres calculs on ne fait intervenir que des additions, des multiplications ou des divisions qui ne portent que sur des termes positifs.

3°) Nous allons maintenant donner une autre façon de résoudre le système $y t = x$, un peu moins classique mais bien adaptée au calcul des probabilités stationnaires.

B.3 Cas limite

1°) On considère le cas particulier $x = 0$ et, quel que soit i , $\sum_{j=1}^m t_{i,j} = 0$ ce qui devient ici :

$$v_1 = u_2, \quad v_m = w_{m-1} \quad \text{et, pour } 1 < i < m, \quad v_i = w_{i-1} + u_{i+1}$$

La famille y_i est alors, à une constante multiplicative près, la "**probabilité stationnaire**" associée à la file $M/M/$. généralisée, les taux d'arrivée u_i et de départ w_i pouvant dépendre de l'état. C'est un cas très particulier où on a toutes les propriétés classiques simultanément (balance locale, forme produit, réversibilité,...). Le sous-espace des solutions y est de dimension un.

2°) Si on impose $y_1 = 1$, y_i est définie par récurrence croissante par $y_i = y_{i-1} u_i / w_{i-1}$.

3°) On peut aussi dire que, si on impose $y_1 = 1$, la séquence $(y_i)_{i \geq 1}$ se détermine "théoriquement" par récurrence croissante sur i en résolvant successivement chaque équation.

4°) Attention : du point de vue numérique, résoudre successivement les équations comme indiqué au 3°), peut introduire des erreurs de calcul nettement supérieures à celles associées à la formule de récurrence donnée au 2°) (à cause des soustractions).

B.4 Cas général

1°) Soit $(b_i)_{1 \leq i \leq m}$ la séquence définie par :

$$b_1 := 1, \quad b_2 := b_1 v_1 w_1^{-1} \text{ et, pour } 2 < i \leq m$$

$$b_i := (b_{i-1} v_{i-1} - b_{i-2} u_{i-1}) w_{i-1}^{-1}$$

(avec, évidemment, $w_{i-1}^{-1} := 1/w_{i-1}$, etc...)

2°) Posons $v'_1 := v_1 - u_2$ et, pour $1 < i < m$, $v'_i := v_i - u_{i+1} - w_{i-1}$ et supposons que l'on ait $u_i \geq 0$, $w_i \geq 0$ et $v'_i \geq 0$ (en B.3 on avait $v'_i = 0$). La construction de b_i donnée au 1°) qui précède implique

$$b_{i+1} w_i - b_i u_{i+1} = b_i v'_i + (b_i w_{i-1} - b_{i-1} u_i)$$

$$\text{et} \quad b_{i+1} w_i = b_i(v'_i + u_{i+1}) + (b_i w_{i-1} - b_{i-1} u_i)$$

On en déduit, en raisonnant par récurrence croissante sur i , que, pour $1 < i \leq m$, b_i et $(b_i w_{i-1} - b_{i-1} u_i)$ sont positifs.

3°) Soit $(c_i)_{1 \leq i \leq m}$ la séquence définie par :

$$c_1 := 0, \quad c_2 := -x_1 w_1^{-1} \text{ et, pour } 2 < i \leq m,$$

$$c_i := (c_{i-1} v_{i-1} - c_{i-2} u_{i-1} - x_{i-1}) w_{i-1}^{-1}$$

4°) On pose :

$$s := (c_{m-1} u_m - c_m v_m + x_m)(-b_{m-1} u_m + b_m v_m)^{-1}$$

ce qui est bien défini si $(-b_{m-1} u_m + b_m v_m)$ est strictement positif (cf. 2°)).

La solution cherchée est alors définie par

$$y_i := s b_i + c_i$$

Plus précisément, quelle que soit la valeur de s , la famille (y_i) définie par $y_i = s b_i + c_i$ satisfait à toutes les équations données en B.1, 2°) sauf celle associée à x_m ; la valeur de s donnée ci-dessus est déterminée par cette dernière équation.

5°) Attention : si $x = 0$, $s = 0$, c'est à dire que la solution obtenue ci-dessus est la solution $y = 0$. L'étude effectuée dans ce paragraphe B.4 est donc distincte et complémentaire de celle effectuée en B.3.

B.5 Algorithme

1°) On garde les hypothèses et notations du paragraphe B.4. Il est

important de noter que l'algorithme associé aux calculs proposés en B.4 est très facile à implémenter. On procède en trois étapes ; l'initialisation et la valeur choisie pour x dépendent de l'étape ; par contre, **à chacune des trois étapes, on effectue le même calcul de z_i** , par récurrence croissante sur i , en prenant $z_i := (z_{i-1} v_{i-1} - z_{i-2} u_{i-1} - x'_{i-1}) w_{i-1}^{-1}$.

2°) A la première étape on prend $x' = 0$ et $z_1 := 1$, ce qui donne la famille (b_i) ; à la deuxième étape on prend $z_1 := 0$ et $x' = x$, ce qui donne la famille (c_i) ; à la fin de la deuxième étape on calcule s ; à la troisième étape on prend $z_1 = s$ et $x' = x$, ce qui donne la famille (y_i) . Notons qu'à chaque pas l'ordinateur n'a à effectuer que des opérations élémentaires : on dit alors que le **schéma de résolution est explicite**.

3°) En procédant comme indiqué précédemment, pour chaque étape fixée, quand on passe de i à $(i+1)$, **il n'est pas nécessaire de garder en mémoire toutes les valeurs de b ou c** : par exemple, il suffit de disposer de b_{i-2} et b_{i-1} pour calculer b_i ; de même, pour calculer s , il suffit de disposer des valeurs b_{m-1} , b_m , c_{m-1} et c_m . En général, les "coefficients" u_i , v_i et w_i sont donnés sous forme de formules simples et ne prennent donc pas de place mémoire. Enfin, seules certaines quantités sont intéressantes et y n'a pas à être mis en mémoire non plus (cf. A.3, 4°).

Par exemple, pour les calculs qui précèdent et si on ne tient pas compte des coefficients, une douzaine de places mémoires suffisent (quelle que soit la valeur de m) : cette remarque prendra tout son intérêt quand les termes u_i , ..., x_i , y_i etc... seront des matrices et non des nombres (cf. la remarque A.3, 4°). Par contre, on ne peut pas faire la même remarque si on utilise une méthode du type factorisation comme indiqué en B.2 ci-dessus.

C Matrices tridiagonales par blocs

C.1 Hypothèses

1°) On adopte les hypothèses et notations données en A.2. De plus, on suppose qu'il existe **une application n définie sur E** , à valeurs dans l'ensemble $\{0, \dots, m\}$, où m est un entier strictement positif, telle que $|n(e) - n(e')| > 1$ implique $a(e, e') = 0$.

On pose $E_k := \{e : e \in E, n(e) = k\}$.

2°) Cette situation est très fréquente dans les systèmes markoviens : **$n(e)$ est, par exemple, le nombre de clients** (dans une station, dans un sous-ensemble du réseau, d'une classe donnée, etc...) et l'hypothèse ci-dessus est satisfaite si ces clients se déplacent individuellement et non par paquets.

En fait, en général, il existe beaucoup de fonctions n pour lesquelles l'hypothèse donnée au 1°) est satisfaite ; au contraire, cette situation (avec beaucoup de fonctions telles que n) est assez rare en analyse numérique classique.

3°) Si on choisit une indexation de E , soit $E = \{e_i, 1 \leq i \leq m'\}$ (avec $m' := \text{card}(E)$), et si cette indexation est telle que $i \leq j$ implique $n(e_i) \leq n(e_j)$, la matrice associée à a pour cette indexation (et donc celle associée à $(a-d)$) est tridiagonale "par blocs".

C.2 Méthode générale de résolution

1°) Avec les hypothèses données ci-dessus, le système $y \cdot t = x$ équivaut exactement au même système formel d'équations que celui proposé en B.1,

2°) sauf que les termes x_i et y_i (pour $1 \leq i \leq m$) sont des **matrices unilignes** et que les termes u_i , v_i et w_i sont des **matrices carrées**.

En fait, on peut même adapter l'étude qui suit au cas où les dimensions des matrices x_i et y_i sont variables, les matrices u_i , v_i et w_i étant alors rectangulaires.

2°) La méthode de résolution proposée en B.3 ou B.4 reste utilisable puisqu'elle ne fait pas intervenir la commutativité. Plus précisément, l'algorithme proposé en B.5 reste opérationnel avec les modifications qui suivent.

3°) Lors de la première étape, au lieu de considérer une seule initialisation $b_1 = 1$, on considère m'' initialisations correspondants à m'' vecteurs de base "positifs" (les m'' matrices unilignes qui constituent la matrice unité). La valeur de m'' sera précisée en C.3, 5°) in fine.

Pour chacune de ces initialisations de z_1 on calcule z_i , par récurrence croissante sur i , comme indiqué en B.5 : notons que **ces calculs peuvent être menés en parallèle** pour les m'' initialisation (cf. la remarque A.3, 5°)).

4°) Considérons le cas où $m'' = \text{card}(E_0)$.

Le recollement des m'' valeurs de z_{m-1} et z_m constitue les matrices b_{m-1} et b_m : en général, la matrice $(-b_{m-1} u_m + b_m v_m)^{-1}$ est inversible. On peut donc, à la fin de la deuxième étape, calculer la matrice $(1 \times m'')$ s et conclure comme en B.5. On a alors **un schéma semi-explicite : plus précisément, seule l'inversion ci-dessus est implicite.**

C.3 Variables auxiliaires

1°) La méthode expliquée ci-dessus consiste donc, en résumé, à introduire m'' inconnues auxiliaires (les m'' coordonnées de la matrice s), à "résoudre" le système global en fonction de ces m'' inconnues, puis à calculer ces inconnues en utilisant les équations non utilisées auparavant. Nous reviendrons sur le calcul de s au paragraphe C.4.

2°) A chaque pas i , on a à résoudre le système

$$-z_{i-1} u_i + z_i v_i - z_{i+1} w_i = x'_i$$

où la matrice inconnue à calculer est la matrice z_{i+1} (les valeurs de z_{i-1} , z_i et x'_i dépendent du stade où en est l'algorithme).

En général, la matrice associée à ce système est elle aussi tridiagonale par blocs : on peut donc procéder comme indiqué ci-dessus mais au niveau "inférieur". Les valeurs des coefficients de la matrice - associée au système à résoudre et qu'en général on n'explicite pas - dépendent du stade où en est l'algorithme. Par contre, la structure du graphe - c'est à dire la place des coefficients non nuls - est presque indépendante de ce stade.

3°) En fait, si on choisit bien l'ordre dans lequel on utilise les fonctions n - comme indiqué en C.1, 1°) - on peut souvent introduire des variables auxiliaires - comme expliqué ci-dessus - à plusieurs niveaux successifs. Par exemple, dans [Sag], on introduit des variables auxiliaires **à trois niveaux successifs**, le niveau 1 étant le niveau global.

4°) Attention toutefois : s'il s'agit d'un calcul de probabilités stationnaires, le premier niveau (analogue à B.3) a des caractéristiques très différentes de celles des autres niveaux (analogues à B.4) : ce point, qui généralise B.3, sera repris en C.4.

5°) Quand on introduit des variables auxiliaires à divers niveaux successifs, à chaque changement de niveau la famille des inconnues auxiliaires fait intervenir un paramètre de moins. Par exemple, dans [Sag], chaque état est naturellement indexé par 4 paramètres ; au niveau 1, les inconnues auxiliaires, si on ne dépassait pas le niveau 1, seraient

indexées par 3 paramètres ; de même, au niveau 2, elles seraient indexées par 2 paramètres.

Comme on peut, dans [Sag] aller jusqu'au niveau 3, la famille des inconnues auxiliaires pour le calcul global est indexée par un seul paramètre : dans ce cas, $m'' = j$ où j est le nombre de valeurs de ce paramètre. Par contre le nombre total d'états est $j k u n$ avec, par exemple, $j = 5$, $k = 40$, $u = 5$ et n quelconque, fini ou non.

6°) Toutefois, **ce gain** considérable en taille mémoire serait **sans intérêt** si cette méthode conduisait à inverser des matrices **très mal conditionnées**, ce qui est naturellement le cas avec ce type de méthode. Tout ce qui précède n'a donc d'intérêt que grâce à la situation spécifique qui va être expliquée au paragraphe D qui suit.

C.4 Retour sur la factorisation

Avant d'expliquer la méthode des convexes, revenons sur la factorisation évoquée en B.2. Cette méthode reste évidemment utilisable quand les termes u_i , v_i et w_i sont des matrices rectangulaires (avec les dimensions adéquates) et non des nombres. Plus précisément, on pose

$$z_j := \text{matrice inverse de } (v_j - w_{j-1} z_{j-1} u_j)$$

etc... (attention, les matrices ci-dessus ne commutent pas, en général).

Par ailleurs, au lieu de définir la séquence z_i pour i croissant et à partir de $i = 1$ on peut la construire à partir de $i = m$ et pour i décroissant (avec les modifications associées évidentes) : ceci est notamment intéressant si les matrices u_i , v_i et w_i sont fixes pour i assez grand : il en est alors de même de z_i . Ce cas a été extensivement étudié par Neuts dans [Neu].

On peut alors montrer (cf. [Neu]) que, pour i assez grand, $x = 0$ et des coefficients fixes à partir d'un certain rang, il existe une matrice carrée positive t telle que

$$y_{i+1} = t y_i$$

et cette matrice a une interprétation probabiliste précise et utilisée depuis très longtemps (par exemple, cf. [Chu]).

Toutefois, attention, y_i n'est pas le vecteur propre associé à la plus grande valeur propre de t : autrement dit, y_i n'est pas la "limite en

direction" de la suite $(t^n y)_{n>0}$ contrairement à ce qui se passe pour les algorithmes proposés plus loin.

Notamment, il semble que, même dans le cas particulier étudié par Neuts, sa démarche ne permet pas de répondre - dans un contexte général - aux questions "théoriques" posées en D.4.

D. La méthode des convexes

D.1 Cadre général

En introduisant des inconnues auxiliaires à divers niveaux on arrive finalement, au niveau 1, (le niveau global) à la situation suivante : le système global à résoudre est $y t := 0$ (avec $t := a-d$) dont on sait qu'il admet une et une seule solution, à une constante multiplicative près, celle-ci étant strictement positive.

On suppose que y est de la forme $y = (y_0, \dots, y_m)$ (m a été introduit en C.1,1°). On suppose qu'il y a un algorithme qui permet de calculer linéairement y_k en fonction de y_0 . De plus, on suppose qu'il existe une "matrice" ($1 \times m$) (une famille d'inconnues) u telle que y_0 soit une fonction linéaire de u : plus précisément, on suppose qu'il y a un algorithme qui permet de calculer y_0 comme fonction linéaire de u .

Formellement, on a donc $y_k = r_k u$ où r_k est une matrice rectangulaire que, en général, on ne cherche pas à expliciter. "Théoriquement", la valeur de u (associée au bord $k = 0$) doit être déterminée par les relations à l'"autre" bord (pour $k = m$).

En fait, pour presque tous les exemples concrets de systèmes markoviens, "ce qui se passe" au bord $k = 0$ est "presque" indépendant de l'autre bord $k = m$ (on peut même avoir m "infini") : **étudier un bord par l'intermédiaire de l'autre bord conduit donc nécessairement à inverser une matrice très mal conditionnée, ce qui est irréaliste.**

D.2 Les convexes $c(k)$

En général, chaque terme de u correspond à la probabilité d'un état : c'est donc un terme positif. De plus, le domaine $c(o)$ auquel u peut appartenir est un convexe parfaitement déterminé. Le plus souvent, $c(o)$ est tout l'octant positif de \mathbb{R}^m . De même, chaque terme de y doit être positif.

Quand on calcule $r_k u = y_k$, les seules valeurs acceptables de u sont

celles pour lesquelles y_k a toutes ses composantes positives. Appelons $c(k)$ le convexe défini par u appartient à $c(k)$ si, quel que soit j , $0 \leq j \leq k$, $r_k u$ a toutes ses coordonnées positives.

Supposons - et ceci est un critère de choix des inconnues - que toutes les valeurs de $c(k)$ soient atteintes si on impose en k des contraintes "artificielles" adéquates (cf. le 6°) de la preuve de E.3).

Si on suppose, ce qui est le cas usuel, que **les contraintes au bord $k = m$, pour m assez grand, influent peu sur la situation du bord $k = 0$** , cela signifie évidemment que **le diamètre de $c(m)$ est petit.**

Au lieu de résoudre les équations au bord $k = m$, on détermine donc le convexe $c(m)$.

Techniquement, il suffit de considérer les points extrémaux des valeurs admissibles de y_m ce qui donne, par l'inversion d'une matrice r' (de dimension $(m \times m)$) les points extrémaux du convexe $c(m)$ (en général, y_k positif implique y_{k-1} positif).

Et si cette matrice r' est mal conditionnée ! c'est que l'on a été trop loin. Les contraintes en m ne sont pas utilisées : on peut donc arrêter l'algorithme pour n'importe quelle valeur de k .

D.3 Mise en oeuvre

1°) Comme pour toutes les mises en oeuvre quand on effectue un calcul scientifique numérique, il y a des choix techniques à effectuer qui sont fonction, non seulement du problème étudié, mais aussi de la valeur des paramètres qui interviennent.

2°) Supposons d'abord que, pour le problème considéré (y compris les valeurs des paramètres), les états "importants" soient les états associés aux "petites" valeurs de $n(e)$. (n a été définie en C.1,1°). Il faut alors procéder comme indiqué précédemment. On arrête les calculs dès que le diamètre de $c(k)$ est assez petit.

3°) Par contre, si les états appartenant à E_0 ont un poids négligeable, il peut être préférable d'initialiser l'algorithme pour $k = j$ et de l'arrêter pour $k = j + j'$. Il faut alors que le domaine $E' := \bigcup_{k=j}^{j+j'} E_k$ contienne les états "importants".

4°) Si, au 3°) qui précède, j' est "trop grand", on peut découper le

domaine "important" E en plusieurs régions E'_i et, pour chaque région E'_i , on procède comme au 3°).

D.4 Convergence

En général (cf. [Alg], [Ast], [BDK], [BoP], [Khm], [Oum], [Pel], [Sag], etc...) le diamètre du convexe $c(k)$ décroît très vite quand k augmente ; cette rapidité est même surprenante (pour ne pas dire plus !).

Ceci introduit les deux questions suivantes qui sont fortement corrélées :

a) Comment prouver que le diamètre de $c(k)$ tend vers zéro quand k tend vers l'infini ?

b) Comment prouver que, lorsqu'un phénomène a une "structure en ligne" comme indiqué en C.1, la situation à un bord dépend peu de celle de l'autre bord (pour m assez grand) ?

Il ne semble pas possible de prouver ce résultat par les méthodes probabilistes classiques : excursions, temps de retours, etc... Une réponse théorique sera donnée au chapitre 7. Pour un cas particulier une réponse simple sera donnée en E.4.

D.5 Un cadre plus précis

Le lecteur mathématicien trouve peut-être que ce paragraphe D est trop "flou" Il peut trouver dans [OPS] des conditions précises qui constituent un cas particulier de D.1 mais qui sont suffisamment générales pour contenir tous les exemples évoqués en D.4.

E Un cas particulier

E.1 Introduction

Nous allons maintenant considérer un **cas particulier** : pour ce paragraphe, l'hypothèse E.3 (i) est **cruciale** mais elle est **très restrictive** et rarement satisfaite. Il semble pourtant intéressant de faire figurer ce paragraphe dans ce cours pour plusieurs raisons.

1°) C'est un exemple précis qui permet d'illustrer l'utilisation de "**contraintes artificielles**" dans la méthode des convexes (cf. le paragraphe D qui précède et le 6°) de la preuve de E.3).

2°) Dans ce cas particulier, il est "facile" de prouver que le **diamètre de $c(k)$ tend vers zéro** quand k tend vers l'infini (cf. E.4) : de ce point de vue, c'est une introduction au chapitre 7.

3°) Cet exemple permet d'illustrer l'utilisation des **matrices tridiagonales par blocs** : au niveau global, on se ramène à une matrice bidiagonale par blocs ; au niveau 2, on se ramène à l'étude d'une matrice triangulaire ou tridiagonale (cf. E.5).

4°) La technique proposée dans ce paragraphe permet de calculer, rapidement, avec précision et pour une faible taille mémoire les probabilités stationnaires de certaines files uniques qui font l'objet d'une littérature surabondante et qui sont importantes au niveau des applications.

5°) Enfin, cette situation étant relativement élémentaire, il est facile de généraliser le théorème E.3 à un contexte où les propriétés spécifiques des matrices a et d sont peu utilisées (cf. E.6). Une partie de ce paragraphe E figure dans [BeP].

E.2 Notations et hypothèses générales

Soit m et m' deux entiers strictement positifs. Soit M (resp. M') l'ensemble des entiers k tels que $0 \leq k \leq m$ (resp. $1 \leq k \leq m'$).

On pose $E := (M' \times M)$: c'est l'ensemble des états. Soit a une fonction positive définie sur $(E \times E)$ telle que, pour tout élément e de E , $a(e, e) = 0$. On suppose que la matrice (généralisée) a est irréductible (cf. 2.B.3).

Le problème est de calculer la famille q des nombres réels $(q_e)_{e \in E}$ qui satisfont aux relations 5A1 et 5A2 (cf. A.2). On sait que cette famille existe, est unique et que chaque terme q_e est strictement positif (cf. 2.D.3).

On définit d comme en A.2 et le système 5A1 équivaut à $q d = q a$.

On suppose que $|h-k| > 1$ implique $a[(i, h), (j, k)] = 0$ c'est à dire que a est tridiagonale par blocs : les notations proposées dans ce paragraphe correspondent à celles proposées en C.1 en posant $n(j, k) := k$. Notamment E_k est l'ensemble des éléments e de E tels que $n(e) = k$.

E.3 Théorème

En plus des hypothèses générales données en E.2 ci-dessus, on suppose qu'il existe deux fonctions positives u et v définies sur E et telles que les deux propriétés suivantes soient satisfaites :

(i) quels que soient i, j et $k, i \in M', j \in M', 1 \leq k \leq m$,
 $a[(i, k-1), (j, k)] = u(i, k) v(j, k)$

(ii) quel que soit $k, 1 \leq k \leq m$, $\sum_{j=1}^{m'} v(j, k) = 1$

Soit q_k la matrice uniligne ($1 \times m'$) dont le terme de la j -ième colonne est défini par :

$$(q_k)_{(1,j)} := q(j, k)$$

Alors, il existe une matrice r' et une famille $(r_k)_{0 \leq k \leq m}$ de matrices ($m' \times m'$) telles que le système $q d = q$ a soit équivalent aux deux conditions suivantes :

(iii) quel que soit $k, 0 \leq k < m$, $q_k = q_{k+1} r_k$

(iv) $q_m = q_{m-1} r'$

La matrice r_0 (resp. r') ne dépend que des valeurs de $a(e, e')$ pour e et e' éléments de $(E_0 \cup E_1)$ (resp. $(E_{m-1} \cup E_m)$). De même, quel que soit $k, 1 \leq k < m$, chaque matrice r_k ne dépend que des valeurs $a(e, e')$ pour e et e' éléments de $(E_{k-1} \cup E_k \cup E_{k+1})$.

Enfin, quel que soit $k, 0 \leq k < m$, la matrice r_k est positive.

Preuve :

1° Pour $0 \leq k < m$, l'équation d'équilibre (relation 5A1) associée à l'état $e = (j, k)$ s'écrit :

$$q(e) \sum_{e'} a(e, e') = \sum_{s=-1}^{+1} w_s(k)$$

en ayant posé, pour $s = -1, s = 0$ et $s = 1$:

$$w(k) := \sum_{i=1}^{m'} q(i, k+s) a[(i, k+s), (j, k)]$$

Evidemment, pour $k = 0$, $w_{-1}(0) = 0$.

2° On pose :

$$t_k := \sum_{x=1}^{m'} \sum_{y=1}^{m'} q(y, k) a[(y, k), (x, k-1)]$$

(pour $k = 0$ on a $t_0 = 0$).

Le système 5A1 implique que, quel que soit k , $0 < k \leq m$, on a :

$$t_k = \sum_{x=1}^{m'} \sum_{y=1}^{m'} q(y, k-1) a[(y, k-1), (x, k)]$$

Cette propriété, classique et élémentaire, figure par exemple dans [Kel] (cf. aussi [Mar] et [Lem] pour son utilisation systématique) : elle sera généralisée au 1°) de la preuve du théorème E.6.

3°) Compte tenu de la propriété (i), la dernière relation ci-dessus implique :

$$t_k = \sum_{x=1}^{m'} v(x, k) \sum_{y=1}^{m'} q(y, k-1) u(y, k-1)$$

soit (cf. (ii)) :

$$t_k = \sum_{y=1}^{m'} q(y, k-1) u(y, k-1) \quad \text{donc}$$

$$w_{-1}(k) = v(j, k) t_k$$

4°) La relation écrite au 1°) devient donc :

$$q(e) \sum_{e'} a(e, e') = v(j, k) t_k + w_0(k) + w_1(k)$$

En utilisant la définition de t_k (début du 2°)), cette relation peut s'écrire matriciellement sous la forme

$$q_k b_k = q_{k+1} c_{k+1}$$

5°) On vérifie facilement, en raisonnant par récurrence croissante sur k , que la matrice b_k est régulière (et donc inversible) sinon le système global des équations 5A1 et 5A2 aurait plusieurs solutions distinctes (pour ce passage, on peut remplacer la matrice c_{k+1} par une matrice c'_{k+1} mieux mélangeante).

6°) On pose alors $r_{k+1} := c_{k+1} b_k^{-1}$. Il faut prouver que r_k est une matrice positive. Pour ce faire, on introduit au "bord k " des contraintes "artificielles" : cette technique a un caractère tout à fait général (cf. par exemple, [Alg]).

Plus précisément, i et k étant fixés, soit a' la fonction définie sur $(E \times E)$ par :

a) $a'(e, e') := a(e, e')$ si e ou e' n'appartient pas à E_k
et, pour $i \neq j$:

b) $a'[(i, k), (j, k)] := (1 + a[(i, k)]) := (1 + a[(i, k), (j, k)]) / z$

c) $a'[(j, k), (i, k)] := z + a[(j, k), (i, k)]$

d) $a'[(j, k), (j, k)] := 0$

Le système global associé à a' est ergodique ; soit q' la solution stationnaire associée. Si r'_{k-1} est défini relativement à a' comme r_{k-1} l'est relativement à a , on a $r'_{k-1} = r_{k-1}$ donc $q'_{k-1} = q'_k r_{k-1}$. Or q'_{k-1} est toujours à termes positifs. Quand z tend vers l'infini, q'_k tend vers un vecteur de base (celui dont la i -ième coordonnée est égale à 1 et toutes les autres sont nulles). Ceci montre que r_{k-1} est une matrice positive.

E.4 Convergence des convexes $c(k)$

On vient de prouver en E.3 que

$$q_0 = q_{k+1} r_k r_{k-1} \dots r_j \dots r_1 r_0$$

En général, les matrices r_k sont "**bien mélangeantes**" c'est à dire que leurs coefficients de contraction (cf. 2.B) sont nettement inférieurs à 1 : q_0 dépend donc peu de q_{k+1} . Plus précisément, quand q_{k+1} parcourt tout l'octant positif de \mathbb{R}^m (ce qui est compatible avec le fait que q_0 soit positif : cf. le 6°) de la preuve de E.3 ci-dessus), q_0 parcourt le convexe $c(k)$ dont le diamètre est inférieur à $\prod_{j=0}^k \mu(r_j)$ où μ est défini comme en 2.B. Le convexe $c(k)$ est l'image de l'octant positif par l'inverse de la matrice $(r_k \dots r_0)$.

E.5 Exemples et remarques

1°) La file PH/M/1. est un cas particulier de la situation proposée en E.3. C'est une file unique pour laquelle la loi de service est exponentielle mais le délai entre deux arrivées suit une loi qu'on peut modéliser par des états fictifs ; tous les taux peuvent dépendre librement de l'état de la file. L'état (j, k) correspond au cas où j est l'état fictif f et il y a k clients dans la

file.

Pour cet exemple, la matrice b_k est "presque" triangulaire (et donc facile à inverser).

2°) La situation considérée en E.3 contient aussi le cas d'une station S munie d'un "buffer" B. La capacité de la station est limitée. Quand cette station S est pleine, les nouveaux clients vont dans le buffer B. Toutes les lois sont exponentielles et tous les taux peuvent dépendre librement de l'état. L'état (j,k) correspond au cas où il y a j clients dans la station S et k dans le buffer B. Dans ce cas, la matrice b_k est tridiagonale (et donc facile à inverser).

Evidemment, le buffer B peut être fictif ; par exemple, il peut modéliser les clients en instance de répétition d'appels à un standard téléphonique.

Pour ces deux exemples, on a un schéma de résolution explicite qui ne fait "presque" pas intervenir de soustractions (cf. A.3,6°). On peut aussi "mélanger" ces deux exemples en considérant une file PH/M/. munie d'un buffer.

3°) La file M/PH/1 à capacité limitée à m est aussi un cas particulier de la situation proposée en E.3 (cf. aussi [MaP]). C'est une file unique pour laquelle le délai entre deux arrivées suit une loi exponentielle et il y a un seul serveur dont la loi de service est modélisable par états fictifs ; la discipline de service est "premier arrivé, premier servi". L'état (j,k) correspond au cas où j est l'état fictif et il y a $(m-k)$ clients dans la file. Dans ce cas, supposer que la capacité est limitée à m a pour seule utilité de pouvoir retrouver le formalisme simple proposé en E.3. En fait, en E.3, on peut supposer que k parcourt un intervalle quelconque de \mathbb{Z} (comme en E.6).

4°) Attention : pour autant, il n'y a pas symétrie entre considérer k en croissant ou en décroissant. Plus précisément, les matrices r_k qui interviennent dans le théorème E.3 sont inversibles mais, évidemment, les matrices r_k^{-1} ne sont pas positives : ce point sera éclairci en 6.B.

5°) Du point de vue technique, puisqu'il s'agit de calculer la matrice inverse de $(r_k \dots r_0)$ on peut, suivant les cas, choisir k en croissant (c'est la présentation proposée en E.3) ou en décroissant.

6°) Pour certains exemples (file PH/M/s notamment), la matrice r_k est fixe pour $k \geq s$ et m est "très grand" (éventuellement "infini"). Dans ce cas, pour $k \geq s$, $q_k = \lambda q_{k+1}$ est fixe "en direction" : c'est le vecteur propre

associé à la plus grande valeur propre de $r = r_k$ (pour $k \geq s$).

7°) Pour plus de détails cf. [Bel] et [BeP].

E.6 Généralisation

Le théorème qui suit généralise E.3 : plus précisément, au lieu de considérer q solution de $q(d-a) = 0$, on considère le système $y h = x$. De plus on suppose que l'ensemble des états peut être infini (on n'utilise donc pas les notations matricielles mais celles-ci sont sous-jacentes). Si E était fini, h serait une matrice tridiagonale par blocs (mais elle ne serait pas nécessairement de la forme $h = d-a$) et h' serait la matrice inverse de h . Là encore, on se ramène à une situation globale "presque" bidiagonale par blocs.

Théorème : Soit K un intervalle (non vide) de l'ensemble \mathbb{Z} des entiers relatifs. Soit K' l'ensemble des entiers relatifs k tels que k et $(k-1)$ appartiennent à K . Soit E un ensemble. Soit $(E_k)_{k \in K}$ une partition de E , chaque ensemble E_k étant non vide. On pose $E(k) := E_k$. Soit a, b, x et y quatre fonctions, réelles ou complexes, définies sur E . Soit h une fonction, réelle ou complexe, définie sur $(E \times E)$. On suppose que les quatre propriétés suivantes sont satisfaites :

$$(5E1) \quad \text{pour tout élément } k \text{ de } K, \quad \sum_{u \in E(k)} b(u) = 1 .$$

$$(5E2) \quad |j-k| > 1, u \text{ élément de } E_j \text{ et } v \text{ élément de } E_k \text{ implique } h(u,v) = 0 .$$

$$(5E3) \quad \text{quel que soit } k \text{ élément de } K', \text{ on a, pour tout couple } (v,u) \text{ élément de } (E_{k-1} \times E_k), h(v,u) = a'(v) b(u) .$$

$$(5E4) \quad \text{la famille } \{ y(e') h(e',e) \}_{(e,e') \in (E \times E)} \text{ est une famille sommable} .$$

$$(5E5) \quad \text{pour tout élément } e \text{ de } E, \text{ on a :}$$

$$\sum_{e' \in E} y(e') h(e',e) = x(e)$$

1°) On introduit les notations suivantes :

$$f(e) := \sum_{e' \in E} h(e,e') ; \quad t'_k := \sum_{v \in E(k)} x(v)$$

$$t_k := \sum_{j > k} \sum_{w \in E(j)} (x-yf)(w)$$

et, si v appartient à $E(k)$:

$$r(v) := \sum_{i=0}^1 \sum_{w \in E(k+i)} h(v, w)$$

$$g(v, u) := h(v, u) - b(u) r(v)$$

$$z(u) := x(u) - b(u)(t_k + t'_k) - \sum_{v \in E(k+1)} y(v) h(v, u)$$

On a alors, pour tout élément k de K' et pour tout élément u de E_k :

$$(5E6) \quad \sum_{v \in E(k)} y(v) g(v, u) = z(u)$$

2°) Soit $\delta_{v,w}$ le symbole de Kronecker, c'est à dire que $\delta_{v,v} := 1$ et $\delta_{v,w} := 0$ pour $v \neq w$. On suppose qu'il existe une fonction h' définie sur $(E \times E)$ telle que, pour tout élément k de K' et pour tout couple (v, w) d'éléments de E_k on ait :

$$(5E7) \quad \sum_{u \in E(k)} h(v, u) h'(u, w) = \delta_{v,w} .$$

On suppose que, pour tout élément k de K' et pour tout élément w de E_k , les familles $\{b(u) h'(u, w)\}_{u \in E(k)}$ et $\{z(u) h'(u, w)\}_{u \in E(k)}$ sont sommables.

On pose $E'(k) := E(k) \times E(k)$.

On suppose que les familles $\{b(u) h'(u, w) r(w)\}_{(u, w) \in E'(k)}$ et $\{z(u) h'(u, w) r(w)\}_{(u, w) \in E'(k)}$ sont sommables.

Pour tout élément k de K' on pose :

$$s'_k = \sum_{(u, w) \in E'(k)} b(u) h'(u, w) r(w)$$

et, pour tout élément w de E_k :

$$b'(w) := \sum_{u \in E(k)} b(u) h'(u, w)$$

On suppose que l'on a $(1-s'_k) \neq 0$. On pose, pour tout élément k de K' :

$$s_k := \sum_{(u, w) \in E'(k)} z(u) h'(u, w) r(w) / (1-s'_k)$$

On a alors, pour tout élément k de K' et pour tout élément w de E_k :

$$(5E8) \quad y(w) = \sum_{u \in E(k)} z(u) h'(u, w) + s_k b'(w)$$

Preuve :

1°) Soit (A, B) une partition de E . On pose :

$$(5E9) \quad q := \sum_{e \in A} \sum_{e' \in B} y(e') h(e', e)$$

On a :

$$q = \sum_{e \in A} \sum_{e' \in E} y(e') h(e', e) - \sum_{e \in A} \sum_{e' \in A} y(e') h(e', e)$$

$$(5E10) \quad q = \sum_{e \in A} x(e) - \sum_{e \in A} \sum_{e' \in A} y(e') h(e', e)$$

2°) Soit k un élément de K' et u un élément de E_k ; on pose :

$$c(i, k) := \sum_{v \in E(k+i)} y(v) h(v, u)$$

Si $(k+1)$ n'appartient pas à K , $c(1, k) = 0$. La relation (5E5) s'écrit alors :

$$(5E11) \quad \sum_{i=-1}^{+1} c(i, k) = x(u)$$

$$3^\circ) \text{ On pose } q_k := \sum_{v \in E(k-1)} y(v) a'(v)$$

On a :

$$(5E12) \quad c(-1, k) = \sum_{v \in E(k-1)} y(v) a'(v) b(u) = b(u) q_k$$

4°) Par ailleurs :

$$q_k = \sum_{v \in E(k-1)} \sum_{u \in E(k)} y(v) a'(v) b(u)$$

On pose $B := \bigcup_{j < k} E_j$ et $A := \bigcup_{j \geq k} E_j$. on a :

$$q_k = \sum_{e \in A} \sum_{e' \in B} y(e') h(e', e)$$

Compte tenu de (5E10) on a donc aussi

$$(5E13) \quad q_k = \sum_{e \in A} x(e) - \sum_{e \in A} \sum_{e' \in A} y(e') h(e', e)$$

5°) Compte tenu de (5E12) et (5E13) la relation (5E11) s'écrit :

$$(5E14) \quad c(o, k) - b(u) \sum_{e \in A} \sum_{e' \in A} y(e') h(e', e) = x(u) - b(u) \sum_{e \in A} x(e) - c(1, k)$$

ce qui équivaut à :

$$(5E15) \quad \sum_{v \in E(k)} y(v) h(v, u) - b(u) \sum_{v \in E(k)} \sum_{i=0}^1 \sum_{w \in E(k+i)} y(v) h(v, w) \\ = x(u) - b(u) \sum_{e \in A} x(e) - \sum_{v \in E(k+1)} y(v) h(v, u) \\ + b(u) \sum_{j > k} \sum_{v \in E(j)} y(v) \sum_{j \geq k} \sum_{w \in E(j)} h(v, w)$$

ce qui équivaut à la relation (5E6) avec les notations indiquées puisque, pour $j > k$ et v élément de $E(j)$, on a :

$$f(v) = \sum_{j \geq k} \sum_{w \in E(j)} h(v, w)$$

6°) La relation (5E6) s'écrit aussi :

$$(5E16) \quad \sum_{v \in E(k)} y(v) h(v, u) = z(u) + b(u) \sum_{v \in E(k)} y(v) r(v)$$

$$\text{On pose } s_k'' := \sum_{v \in E(k)} y(v) r(v) .$$

En multipliant (5E16) par $h'(u, w)$, en sommant sur u et en utilisant la relation (5E7), on obtient :

$$(5E17) \quad y(w) = \sum_{u \in E(k)} z(u) h'(u, w) + b'(w) s_k''$$

7°) En multipliant (5E17) par $r(w)$ et en sommant sur w , on obtient :

$$(5E18) \quad s_k'' = \sum_{w \in E(k)} y(w) r(w) = \sum_{(u, w) \in E'(k)} z(u) h'(u, w) r(w)$$

$$+ s_k'' \sum_{w \in E(k)} b'(w) r(w)$$

$$\text{Or, } \sum_{w \in E(k)} b'(w) r(w) = s_k' .$$

La relation (5E18) implique donc $s_k'' = s_k'$ (pour $(1 - s_k') \neq 0$). La relation (5E17) est alors équivalente à la relation (5E8).

E.7 Conclusion

De ce chapitre 5, il faut retenir essentiellement deux idées : d'une part, calculer des probabilités stationnaires c'est résoudre un système linéaire et on peut utiliser l'arsenal de l'analyse numérique classique ; d'autre part ce système linéaire a des propriétés spécifiques (liées notamment à la "positivité") et il y a souvent lieu de choisir une méthode qui tient compte de ces propriétés.

Les chapitres 6 et 7 qui suivent complètent ce qui précède dans deux directions distinctes et parfois complémentaires.

Notons aussi qu'il y a une littérature abondante sur "l'analyse en valeurs moyennes" et/ou "l'analyse opérationnelle". Ces travaux se situent à mi-chemin entre l'étude théorique générale des systèmes markoviens discrets (cf. les chapitres 1 et 2) et la présentation algébrique comme proposé dans ce chapitre. Nous ne résumerons pas ces travaux dont l'intérêt est essentiellement technique et/ou "pédagogique".

Chapitre 6

Splitting, chaîne incluse et exemples**A Splitting****A.1 Introduction**

Il y a deux aspects dans ce chapitre 6. D'une part, l'un des objectifs en est d'expliquer la méthode de la **chaîne incluse** (cf. le paragraphe C). Le lecteur probabiliste ne comprendrait pas que cette méthode, considérée comme une méthode de base, ne soit pas évoquée dans un cours qui se propose de présenter toutes les techniques effectivement opérationnelles.

Autant que l'auteur ait pu en juger, c'est essentiellement dans le cas "avec pivot" (cf. le paragraphe B) que cette méthode est techniquement efficace.

D'autre part, le point essentiel de ce chapitre relève plus de la méthodologie que des résultats techniques.

Les paragraphes A et B portent essentiellement sur l'utilisation des **splittings**, méthode désormais classique en analyse numérique. L'interprétation "probabiliste" de cette technique est donnée au paragraphe C. Cette interprétation peut, par exemple, aider à trouver de "bons" splittings.

Le cas des "**points absorbants**", crucial en fiabilité, est abordé au paragraphe D. Au paragraphe E, on montre que la discrétisation des équations aux dérivées partielles (en abrégé E.D.P.) linéaires peut conduire à l'étude d'un système linéaire de la forme $y d - y a = x$, la somme des colonnes de la matrice $(d-a)$ étant nulle. On note au paragraphe F que ce même système apparaît dans l'étude des matrices de Léontief. Un essai de conclusion est donné au paragraphe G.

A.2 Théorème

Soit a, b, c et d quatre matrices carrées $(n \times n)$. On suppose que $a = b+c$ et que la matrice $(d-c)$ est inversible. On se propose de résoudre le système suivant :

$$(6A1) \quad y d - y a = x$$

On pose :

$$(6A2) \quad r := (d-c)^{-1} b$$

et

$$(6A3) \quad z = y(d-c)$$

Le système 6A1 équivaut alors à :

$$(6A4) \quad z = z r + x$$

vérification

Notons d'abord que 6A3 équivaut à $y = z(d-c)^{-1}$ donc y est déterminé si et seulement si z est déterminé (ceci ne signifie pas que y est unique). Or 6A1 équivaut à $y(d-c) = y b + x$ et donc à $z = y b + x$. Mais 6A2 implique que $b = (d-c)r$ et donc que $y b = y(d-c) r = z r$. Finalement, 6A1 équivaut à $z = z r + x$.

A.3 Remarques

1°) Le lecteur trouve peut-être humoristique d'appeler théorème la transformation élémentaire proposée ci-dessus. Certes, l'essentiel reste à faire, c'est à dire **choisir une bonne décomposition** $a = b+c$. Toutefois, la transformation ci-dessus est très profonde : elle permet, en un certain sens, de "séparer" l'étude de c de celle de b .

2°) Cette transformation est utilisée depuis plus d'un demi-siècle en probabilité dans le cadre de la "chaîne incluse" (cf. le paragraphe C qui suit). Le lecteur probabiliste est peut-être un peu surpris (voire déçu !) de constater que le coeur "théorique" de la chaîne incluse se résume à une transformation matricielle élémentaire : en fait, ceci en augmente considérablement le niveau de généralité et, surtout, le champ d'application (cf. les remarques 6°) et 7°) qui suivent).

3°) Cette transformation est aussi fréquemment utilisée en analyse numérique sous le nom de splitting. Plus précisément, posons $f := d-a$, $m := d-c$ et $n := b$; on dit que $f = m-n$ est un splitting régulier (cf. par exemple, [GoM], 4.40) de la matrice f si la matrice m est inversible et si m^{-1} et n sont des matrices positives.

4°) Aussi bien en probabilité qu'en analyse numérique classique, on a souvent la situation suivante : les matrices d , b et c sont positives, la matrice d est diagonale et les matrices b et c ont toutes leurs termes de la diagonale qui sont nuls. De plus, soit s la matrice unicolonne dont tous les termes sont égaux à 1. On a souvent $d s \geq a s$ (c'est à dire que chaque terme de la matrice $d s$ est supérieur ou égal au terme associé de la matrice $a s$) : on dit alors que $(d-a)$ est à diagonale dominante.

Notons que, dans le cas du système introduit en 2.D, on a $d s = a s$, propriété qui sera intensivement utilisée au chapitre 7 qui suit.

5°) Le cas le plus important (aussi bien en probabilité qu'en analyse

numérique classique) est le cas où le rayon spectral (c'est à dire le maximum des modules des valeurs propres) de la matrice $(d^{-1} c)$ est strictement inférieur à 1, la matrice d étant inversible. Dans ce cas, la série matricielle de terme général $(d^{-1} c)^k$ est convergente (la convergence est même "géométrique"), la matrice $(d-c)$ est inversible et on a :

$$(d-c)^{-1} = \left\{ \sum_{k=0}^{\infty} (d^{-1} c)^k \right\} d^{-1}$$

Evidemment, dans ce cas, si d^{-1} et c sont des matrices positives, il en est de même de $(d-c)^{-1}$. Notons que, en général, calculer la somme des termes de la série donnée ci-dessus n'est pas la "bonne" méthode pour calculer $(d-c)^{-1}$; toutefois elle est au coeur de la théorie du potentiel (cf., par exemple, [Rev]).

6°) Nous verrons au paragraphe C qui suit que l'étude du système 6A2 (c'est à dire, essentiellement, l'inversion de la matrice $(d-c)$) a une interprétation probabiliste précise si les matrices d et c sont positives, d étant diagonale et c ayant tous les termes de sa diagonale nuls.

Cette interprétation vaut donc même si le problème initial est un problème d'analyse numérique classique (avec $x \neq 0$) n'ayant pas a priori d'interprétation probabiliste.

Certes, au niveau des démonstrations, il est plus facile de raisonner en termes matriciels que d'utiliser la chaîne incluse. Par contre, des interprétations probabilistes analogues à celles de la chaîne incluse peuvent aider à découvrir de "bons" splittings.

7°) De plus, **et ce point nous semble fondamental**, la proposition qui suit montre que, pour inverser la matrice $(d-c)$, on peut se ramener à un système de la forme $q'a' = q'd'$ avec $a's = d's$, ce qui permet notamment d'utiliser tout ce qui sera vu au chapitre 7.

8°) Enfin si la matrice $(d-c)^{-1}$ est positive (cf. le 5°) qui précède) et si $d s \geq a s$ (cf. le 4°)) on a $(d-c)s \geq b s$ et donc $s \geq r s$, ce qui est l'analogie de la propriété 6A1 pour le système 6A4.

A.4 Inversion de $(d-c)$

Proposition : Soit $(d-c)$ une matrice carrée $(n \times n)$. Soit s la matrice unicolonne $(n \times 1)$ dont tous les termes sont égaux à 1. On pose $f := (d-c) s$. Soit $\Phi(i, f)$ la matrice $(n \times n)$ dont toutes les colonnes sont nulles sauf la i -ième qui est égale à f .

Pour tout entier i , $1 \leq i \leq n$, soit r_i une matrice uni-ligne telle que l'on

ait à la fois $r_i(d-c-\Phi(i,f)) = 0$ et $r_i f = 1$. Soit r la matrice carrée $(n \times n)$ dont la i -ième ligne est r_i . On a alors $r = (d-c)^{-1}$. Réciproquement, si $r = (d-c)^{-1}$, la i -ième ligne r_i de r est telle que $r_i[d-c-\Phi(i,f)] = 0$.

Vérification

Notons d'abord que la matrice $(d-c)$ est quelconque, mais on n'a pas voulu changer de notation par rapport au théorème précédent. Par ailleurs, la matrice $r_i(d-c) = r_i \Phi(i,f)$ est une matrice uniligne dont tous les termes sont nuls sauf le i -ième (par définition de $\Phi(i,f)$). De plus, ce i -ième terme vaut 1 puisque $r_i f = 1$. On a donc $r(d-c) = u$ où u est la matrice unité. Réciproquement, si $r(d-c) = u$, $r_i f = r_i(d-c)s = 1$ d'où $r_i[d-c-\Phi(i,f)] = 0$.

Remarque :

La condition $r_i f = 1$ est une condition de normalisation. L'important est donc de trouver (quel que soit i) une matrice uniligne r'_i telle que $r'_i f \neq 0$ et telle que $r'_i(d-c-\Phi(i,f)) = 0$. Cette nouvelle écriture est importante parce que (par construction de $\Phi(i,f)$) la matrice $(d-c-\Phi(i,f))$ est telle que $(d-c-\Phi(i,f))s = 0$.

Par conséquent, si f est positive, si d est une matrice diagonale, si tous les termes de la diagonale de c sont nuls et si d et c sont des matrices positives (c'est à dire si $(d-c)$ est une L-matrice à diagonale dominante), **la matrice $(d-c-\Phi(i,f))$ satisfait à toutes les propriétés qui seront intensivement utilisées au chapitre 7 qui suit.**

A.5 Cas tridiagonal par blocs

Supposons que l'on ait à résoudre le système $y d = y a + x$ où a et d sont des matrices (positives), d étant diagonale (et tous les termes de la diagonale de a étant nuls). On suppose que d et a sont des matrices $(n \times n)$ et on pose $E := \{i : 1 \leq i \leq n\}$. Supposons, de plus, qu'il existe une application m définie sur E et à valeurs dans \mathbb{N} telle que $|m(e) - m(e')| > 1$ implique $a_{e,e'} = 0$. Par exemple, s'il s'agit d'un problème de files d'attente, $m(e)$ est le nombre de clients (d'une certaine classe) dans une station S .

Il est alors parfois opportun de **définir le splitting** $d-a = (d-c)-b$ (soit $a = b+c$) **de la façon suivante :**

$$b_{i,j} := a_{i,j} \text{ si } m(j) < m(i)$$

sinon $b_{ij} := 0$
 $c = a - b$.

Autrement dit, dans le cas d'un problème de files d'attente, b_{ij} est associé aux transitions qui correspondent à un départ de S d'un client de la classe considérée (évidemment on peut considérer les arrivées au lieu des départs).

Les matrices b et c sont "triangulaires par blocs" ce qui simplifie notablement les calculs.

B Cas avec pivot

B.1 Introduction

La terminologie "avec pivot" utilisée dans ce paragraphe renvoie à [Pel.5] (cf. 1.F.4) et non pas à la notion de pivot telle qu'on l'utilise dans la méthode de Gauss, même si ces deux terminologies présentent une certaine analogie.

Plus précisément, considérons un système de files d'attente.. Soit S une station de ce système munie d'un serveur unique dont la durée de service soit modélisable par états fictifs avec pivot comme indiqué en 1.F.4. Dans ce cas, le passage de l'état e à l'état e' ne peut correspondre à un départ que si l'état fictif (du serveur de S) associé à e' est un état pivot. Autrement dit, b_{ij} est nul pour tous les états j tels que l'état fictif associé n'est pas l'état pivot. L'objet de ce paragraphe B est de formaliser ce cas particulier auquel on peut presque toujours se ramener quand on étudie la probabilité stationnaire associée à un réseau de files d'attente.

B.2 Théorème

On utilise la notation par blocs (cf. 5.A.1). On garde les hypothèses et notations proposées dans le théorème A.2, notamment $a = b + c$. De plus, on suppose que $y = (y', y'')$, $x = (x', x'')$, $d = \begin{pmatrix} d' & 0 \\ 0 & d'' \end{pmatrix}$, $b = \begin{pmatrix} b' & 0 \\ b'' & 0 \end{pmatrix}$ et $c = \begin{pmatrix} t & u \\ v & w \end{pmatrix}$.

On suppose que la matrice $(d'' - w)$ est inversible. Comme en A.2, on se propose de résoudre le système 6A1 que l'on rappelle :

$$(6B1) \quad y d - y a = x$$

On pose :

$$(6B5) \quad r' := (d'' - w)^{-1}$$

et

$$(6B6) \quad y' := (y' u + x'') r'$$

Le système 6B1 équivaut alors à :

$$(6B7) \quad y'(d' - b' - t - u r' b'' - u r' v) = x' + x'' r' (b'' + v)$$

Comme en A.2, la vérification de ce théorème est immédiate. Le point important est de penser à aborder le problème par cette méthode.

B.3 Remarques

1°) La simplification par rapport au cadre général A.2 tient évidemment au fait que le système 6B7 - qui est le système le plus difficile à résoudre car c'est lui qui fait intervenir le plus grand nombre de paramètres - ne fait intervenir que y' et non pas y'' . Cette méthode est donc intéressante si le nombre de termes de y' est beaucoup plus petit que le nombre de termes de y'' : ceci est évidemment le cas pour les réseaux de files d'attente avec pivot comme expliqué en B.1.

2°) Evidemment, la proposition A.3 relative à l'inversion de $(d-c)$ reste utilisable dans le cadre de ce paragraphe B.

B.4 Cas tridiagonal par blocs

Dans le cas de la "chaîne incluse", on a une situation "tridiagonale par blocs" (cf. A.5). Pour la commodité du lecteur nous allons expliciter les hypothèses minimales correspondant au cas où on a à la fois B.2 et A.5.

Comme dans le théorème A.2, on considère deux matrices carrées $(n \times n)$ a et d avec d matrice diagonale. On se propose de résoudre le système 6B1, c'est à dire :

$$(6B1) \quad y d - y a = x$$

On pose $E := \{i : 1 \leq i \leq n\}$ et on suppose qu'il existe une application m définie sur E et à valeurs dans \mathbb{N} telle que $|m(e) - m(e')| > 1$ implique $a_{e,e'} = 0$. On pose alors (comme en A.5) :

$$b_{i,j} := a_{i,j} \quad \text{si } m(j) < m(i)$$

$$b_{i,j} := 0 \quad \text{si } m(j) \geq m(i)$$

$$c := a - b$$

De plus, on suppose qu'il existe une partie E' de E (les états "pivot") telle que $m(j) < m(i)$ et j non élément de E' implique $a_{ij} = 0$. Les hypothèses données en B.2 sont alors satisfaites, y' étant associée à E' et y'' étant associée à $E'' := E \setminus E'$. De plus, la matrice w est triangulaire par blocs ce qui facilite la résolution du système 6B5.

C. Chaîne incluse

C.1 Introduction

On a déjà noté à plusieurs reprises que les résultats donnés précédemment dans ce chapitre ont une interprétation probabiliste : celle-ci va être explicitée dans ce paragraphe C. Ces résultats admettent une démonstration matricielle élémentaire : il ne semble donc pas indispensable d'en donner, par ailleurs une démonstration probabiliste rigoureuse, qui, de toutes façons, nécessite des hypothèses de positivité très restrictives. Ce paragraphe C sera donc essentiellement heuristique.

Plus précisément, le but de ce paragraphe C est de montrer comment **l'interprétation probabiliste permet de découvrir des propriétés matricielles**. Nous allons commencer par expliquer l'interprétation associée à la proposition A.4 ce qui introduit cette technique à partir d'un exemple simple.

C.2 Inversion de (d-c)

1°) On considère donc quatre matrices carrées ($n \times n$) positives a, b, c et d , avec $a = b + c$. On suppose que d est diagonale et que tous les termes de la diagonale de a (et donc de b et c) sont nuls. Soit s la matrice uni-colonne dont tous les termes sont égaux à 1. On suppose que $d s = a s$ (comme en 2.D.1).

2°) Dans tout ce qui suit, l'entier $i, 1 \leq i \leq n$, est fixé.

Soit u la matrice unité ($n \times n$) et u_i sa i -ième ligne. Soit $r = (d - c)^{-1}$. La i -ième ligne de r est telle que $r_i(d - c) = u_i$; nous allons interpréter cette relation de la façon suivante :

Soit $E := \{j : 1 \leq j \leq n\}$. Soit $E' := E \cup \{0\}$, c'est à dire que l'on adjoint à l'ensemble E des états un état "extérieur" que l'on appelle 0. Soit a' la matrice d'évolution associée à E' et définie de la façon suivante :

si j et k appartiennent à E , $a'_{j,k} := c_{j,k}$

$$a' := \begin{pmatrix} 0 & u(i) \\ b s & c \end{pmatrix}$$

$$a'_{0,k} := \delta_{i,k}, \quad a'_{k,0} := (b\ s)_k \quad \text{et} \quad a'_{j,j} := 0.$$

Soit d' la matrice diagonale (associée à a') telle que $d's' = a's'$ où s' est la matrice unicolonne $((n+1) \times 1)$ dont tous les termes sont égaux à 1. On note que $d'_{0,0} := 1$ et que, pour $j \neq 0$, $d'_{j,j} = d_{j,j}$.

3°) L'évolution associée à a' peut être interprétée comme suit : pour a , sur E , il y a deux types de transitions (celles associées à b et celles associées à c). Les transitions associées à c restent les mêmes pour a' : par contre, les transitions associées à b sont transformées en des transitions vers le point "extérieur" 0. Le "taux de service" en ce point 0 vaut 1 et la seule transition autorisée quand on quitte l'état 0 est d'aller dans l'état i (i est fixé).

4°) Soit $r'_i := (1, r_i)$: r'_i est donc une matrice $((n+1) \times 1)$.

On a $r_i(d-c) = u_i$ si et seulement si $r'_i(d'-c') = 0$ (rappelons que i est fixé).

Autrement dit, à la constante de normalisation près, r_i est, en restriction à E , la probabilité stationnaire associée à a' .

5°) Il est évident intuitivement que, à la constante de normalisation près, cette probabilité stationnaire, en restriction à E , ne change pas quand on modifie le taux de service en 0, notamment quand on le fait tendre vers l'infini (on ne reste pas en 0). Si on formalise ceci, on obtient exactement la proposition A.4.

C.3 Probabilité stationnaire pour une trajectoire

Pour toute la suite de ce paragraphe C, on va utiliser intensivement les notations introduites aux paragraphes 2.E et 2.F. Notamment, on considère une trajectoire f au sens défini en 2.F.2. Soit E l'ensemble des états de f .

Soit j un élément de E . Soit $x_j(t)$ la fonction (markovienne) définie par $x_j(t) = 1$ si $f(t) = j$, sinon $x_j(t) = 0$. On définit $m(s, t, x_j)$ comme en 2.F.2 et on étudie la limite de $m(s, t, x_j)$ quand t tend vers l'infini.

En général, cette limite $m'(x_j)$ existe et elle ne dépend ni de s , ni de la trajectoire considérée, mais seulement de la loi d'évolution du processus considéré (cf. 2.F.3). S'il en est ainsi (cf. 2.E.2), $m'(x_j)$ est la probabilité stationnaire d'être dans l'état j . Si la loi d'évolution de la trajectoire considérée f est régie par la matrice irréductible a (cf. 2.D), on a $m'(x_j) = q_j$ où q est la solution normalisée du système $q d = q a$ (où d est la matrice

diagonale associée à a).

En langage imagé mais, nous semble-t-il, évocateur, on peut dire (cf. [Pel.6]) que $m'(x_j)$ est le pourcentage de la "place dans le temps" prise par les cas où l'état est j .

C.4 Transition H

1°) On considère toujours la trajectoire fixée f . De plus, on se donne une partition (H, H') de $(E \times E)$. On dira qu'une transition (e, e') (c'est à dire de e à e') est de type H (resp. H') si (e, e') appartient à H (resp. H'). Observons la trajectoire f à l'instant t ; soit u le maximum des éléments v de $(0, t)$ tels que il existe (e, e') élément de H avec $f(v^-) = e$ et $f(v^+) = e'$; on dira que u est l'instant de la dernière transition de type H (avant t) et que la dernière transition de type H est (e, e') . Si l'ensemble ci-dessus est vide, on pose $u = 0$ (par convention).

2°) Soit (i, j) un couple d'éléments de E . On considère les fonctions suivantes :

$y_i(t) = 1$ si, lors de la dernière transition (e, e') de type H on avait $e' = i$; on pose $y_i(t) = 0$ dans les autres cas ; s'il n'y a pas eu de transition de type H avant t on pose $y_i(t) = 0$.

On pose :

$$z_{ij}(t) = 1 \text{ si } f(t) = j \text{ et si } y_i(t) = 1$$

(autrement dit $z_{ij}(t) = x_j(t) y_i(t)$ où $x_j(t)$ est définie comme en C.3).

On a évidemment $m(s, t, x_j) = \sum_i m(s, t, z_{ij})$ et $m(s, t, y_i) = \sum_j m(s, t, z_{ij})$.

3°) Notons que y_i n'est pas "presque markovienne" au sens donné en 2.F.3 (car y_i peut dépendre d'un passé "lointain") mais on peut rendre y_i markovienne en mettant en mémoire l'état e' immédiatement après la dernière transition de type H , c'est à dire en grossissant l'ensemble des états du processus en sorte qu'il prenne en compte cette information.

Attention : du point de vue théorique, la construction ci-dessus dépend de H ce qui revient, en fait, à considérer une nouvelle matrice d'évolution. Ceci montre que certaines preuves, apparemment savantes, qui utilisent un langage sophistiqué ne sont valables quant au fond que si elles tiennent compte, d'une façon ou d'une autre, de cette remarque.

4°) En général, les quantités $m(s, t, x_j)$, $m(s, t, y_i)$ et $m(s, t, z_{i,j})$ ont une limite quand t tend vers l'infini et cette limite m' ne dépend ni de s ni de la trajectoire considérée (cf. 2.F.3).

Dans ce qui suit, on suppose qu'il en est ainsi.

5°) On pose $r'_{i,j} := m'(z_{i,j}) / m'(y_i)$.

En langage imagé, $r'_{i,j}$ est le pourcentage de place dans le temps prise par les cas où l'état est j en restriction aux cas pour lesquels la dernière transition (e, e') de type H est telle que $e' = i$.

6°) Donnons une présentation "plus expérimentale" de cette expression. On suppose que l'on dispose de trois chronomètres X, Y et Z. En fait, pour chaque valeur de j (resp. i , (i,j)), il faut un chronomètre X (resp. Y, Z).

On met X en marche quand la trajectoire est dans l'état j : sinon on arrête X.

On met Y en marche dès qu'il y a une transition (e, e') de type H telle que $e' = i$ et on arrête Y dès qu'il y a une transition (e, e') de type H telle que $e' \neq i$.

Z marche si et seulement si X et Y sont en marche.

$t m(s, t, y_i)$ (resp. $t m(s, t, z_{i,j})$) est le temps (cumulé), compris entre s et $s+t$, durant lequel le chronomètre Y (resp. Z) marche.

7°) Pour les chronomètres (ou "les observateurs") Y et Z, tout se passe comme si le temps pendant lequel Y ne marche pas était "effacé". Autrement dit, pour Y et Z, dès qu'intervient une transition (e, e') de type H avec $e' \neq i$, on arrête toutes les horloges et on les relance dès qu'intervient une transition (e, e') de type H telle que $e' = i$. **Pour Y et Z, tout se passe donc "comme si", dès qu'intervient une transition de type H, le processus revenait aussitôt dans l'état i .**

La **matrice r'** construite au 5°) est donc (avec les hypothèses indiquées au 4°) - aux constantes de normalisation près - **la même que la matrice r définie dans la proposition A.4**, la matrice b étant associée aux transitions de type H (cf. C.2). Plus précisément :

$$r'_{i,j} = r_{i,j} / (\sum_k r_{i,k})$$

C.5 Processus inclus

1°) On considère toujours la trajectoire fixée f , les ensembles H et H' et les "chronomètres" X, Y et Z introduits au 6°) du paragraphe C.4 ci-dessus. On se propose maintenant d'étudier Y, c'est à dire $m(s, t, y_i)$.

2°) Pour cela, considérons la fonction g définie sur \mathbb{R}^+ par $g(t) = k$ si et seulement si lors de la dernière transition (e, e') de type H antérieure à t on avait $e' = k$. A chaque trajectoire "brute" f on associe ainsi une autre trajectoire "brute" g .

Par conséquent, si on a bien défini le processus initial X par trajectoires, cette transformation associe à X un autre processus Y que l'on appelle le "processus inclus dans X " associé aux transitions de type H .

3°) La limite de $m(s, t, y_i)$ quand t tend vers l'infini ne dépend pas, en général, de chaque trajectoire de Y mais uniquement de la loi d'évolution de Y . Dans ce cas, cette limite $m'(y_i)$ de $m(s, t, y_i)$ est la probabilité stationnaire pour le processus Y d'être dans l'état i . Cette limite est donc la solution q' d'un système linéaire de la forme $q'd' = q'a'$ (au sens donné en 2.D.1). La matrice a' est associée à l'évolution de Y . Evidemment, on détermine la matrice a' à l'aide de la matrice r définie précédemment.

4°) On a :

$$m(s, t, x_j) = \sum_i m(s, t, z_{i,j})$$

Si ces expressions ont des limites quand t tend vers l'infini, on a donc aussi :

$$m'(x_j) = \sum_i m'(z_{i,j})$$

Or

$$m'(z_{i,j}) = m'(y_i) \{ m'(z_{i,j}) / m'(y_i) \}$$

On a vu en C.4 comment calculer

$$m'(z_{i,j}) / m'(y_i) = r'_{i,j}$$

On a vu au 3°) qui précède comment calculer $m'(y_i)$ ce qui donne la probabilité stationnaire cherchée $m'(x_j)$.

Si on formalise ceci avec les simplifications évidentes, on obtient exactement le théorème A.2 que, évidemment !, il est plus simple de vérifier directement. Ce point sera repris en 8.F.4.

C.6 Chaîne incluse

1°) La présentation proposée en C.4 ci-dessus n'est pas la seule possible. Notamment on fait souvent intervenir ce qu'on appelle usuellement la chaîne incluse.

On considère toujours la trajectoire fixée f ainsi que H et H' . La

trajectoire f est associée à la suite $(e_k, s_k)_{k \geq 0}$ comme cela est expliqué en 2.F.2. Soit u_i la fonction "presque markovienne" définie sur \mathbb{N} par $u_i(k) = 1$ si $e_k = i$ et si (e_{k-1}, e_k) appartient à H , sinon $u_i(k) = 0$. On définit $n(s, t, u_i)$ comme en 2.F.2. De même, soit v la fonction définie sur \mathbb{N} par $v(k) = 1$ si (e_{k-1}, e_k) appartient à H , sinon $v(k) = 0$.

En général, quand t tend vers l'infini, $n(s, t, u_i)$ et $n(s, t, v)$ ont des limites $n'(u_i)$ et $n'(v)$ qui ne dépendent ni de s ni de la trajectoire considérée.

2°) Le 1°) peut être interprété de la façon suivante : on effectue une observation à temps discret (et non pas à temps continu comme au paragraphe C.4). Plus précisément, on observe la trajectoire aux instants (les "temps d'arrêt") qui suivent immédiatement une transition de type H . D'une part on compte le nombre de telles transitions (ce qui donne $n(s, t, v)$) ; d'autre part, on compte le nombre de telles transitions (e_{k-1}, e_k) pour lesquelles $e_k = i$ (ce qui donne $n(s, t, u_i)$).

Le rapport $w_i := n'(u_i)/n'(v)$ est donc, relativement à cette observation, la probabilité stationnaire d'observer l'état i .

3°) La chaîne de Markov (cf. 1.C) associée à l'observation expliquée ci-dessus est appelée la "chaîne incluse" associée aux transitions de type H . La matrice $w := (w_i)_{1 \leq i \leq n}$ est la probabilité stationnaire associée à cette chaîne. C'est "donc" (avec les réserves théoriques évoquées précédemment) la solution normalisée du système $q'' = q''a''$ où a'' est la matrice stochastique (cf. 1.C.3) définie par :

a''_{ij} est la probabilité pour que l'état soit j juste après la $(k+1)$ -ième transition de type H sachant que l'état était i juste après la k -ième transition de type H .

Evidemment on pourrait définir a''_{ij} comme limite de valeur moyenne d'une variable comptage (comme on l'a fait précédemment pour u_i).

4°) La matrice a'' est aussi la matrice stochastique associée au processus inclus introduit en C.5. On a donc $a'' = (d')^{-1} a'$ où a' et d' ont été définis en C.5, 3°) (cf. 2.D.3 et 3.E.1).

5°) L'étude de la chaîne incluse sera reprise en 8.F.4.

D Points absorbants

D.1 Introduction

Soit X un processus qui évolue à temps continu et qui admet l'ensemble fini E comme ensemble d'états. On pose $a(u,u) := 0$ et, pour

$$u \neq v, \quad a(u,v) := \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba} [X_{t+h} = v \mid X_t = u] \right\} \quad (\text{cf. 1.C}).$$

$$\text{On pose } d_{u,v} := \delta_{u,v} \sum_w a(u,w) \quad (\text{cf. 2.D.1})$$

On dit que u est un point absorbant (isolé) si $\sum_v a(u,v) := 0$. On dira que U partie de E est absorbante si $\sum_{v \notin U} \left(\sum_{u \in U} a(u,v) \right) = 0$.

S'il y a au moins un point absorbant u , la matrice a ne peut pas être irréductible (puisqu'on ne peut pas "quitter" l'état u). S'il y a au moins deux points absorbants, le système $q \cdot a = q \cdot d$ admet plusieurs solutions normalisées distinctes ; celles-ci "chargent" les parties absorbantes deux à deux disjointes indépendamment les unes des autres (mais en fonction de la loi initiale).

Sous cet aspect, le cas où il y a des points absorbants (ou des parties absorbantes autres que E tout entier) semble totalement nouveau par rapport aux hypothèses adoptées dans les chapitres précédents. L'évolution du phénomène peut avoir une fin puisqu'elle s'arrête dès qu'on a atteint un état absorbant.

D.2 Probabilité stationnaire

1°) Reprenons l'exemple introduit en 3.D.2. Si on se limite à l'étude d'une partie, on arrête cette étude dès que le joueur est ruiné (cas $Y = 0$) ou comblé (cas $Y = 2^{m+1}$) : tous les états associés à ces valeurs de Y sont des états absorbants.

Evidemment, l'étude stationnaire n'a pas d'intérêt si elle ne prend en compte qu'une seule partie. Par contre il est intéressant de connaître la probabilité pour un joueur de finir ruiné ou, au contraire, comblé.

2°) Si on procède par simulation, la technique proposée en 3.D.2 consiste tout simplement à repartir de l'état initial w (soit $X'_1 = 0$ et $Y'_1 = 2^{n+1}$) dès qu'on atteint un état absorbant (sans rester dans cet état absorbant).

3°) Considérons alors le cadre général proposé en D.1 ci-dessus,

chaque partie absorbante étant réduite à un point isolé et soit K l'ensemble des états absorbants. Considérons le cas où l'état initial w est imposé. Du point de vue mathématique, le 2°) qui précède revient à étudier la probabilité stationnaire associée à (E', a') où E' et a' sont définis comme suit.

Evidemment on suppose que w n'appartient pas à K et on pose $E' := E \setminus K$. Soit (u, v) un couple d'éléments de E' . On pose :

$$a'(u, v) := a(u, v) \text{ si } v \neq w \text{ et } v \text{ n'appartient pas à } K$$

$$a'(u, w) := a(u, w) + \sum_{v \in K} a(u, v) \text{ si } u \neq w$$

$$a'(w, w) := 0$$

On suppose que a' est irréductible. Soit d' la matrice diagonale associée à a' et q' la solution normalisée (cf. 2.D.1) de $q'd' = q'a'$. Pour tout élément e de E' , q'_e est la probabilité stationnaire d'être dans l'état e .

4°) Toutes les quantités "intéressantes" se déduisent facilement de la famille q' . Par exemple, posons, pour tout élément u de K :

$$t'_u := \sum_{e \in E'} q'(e) a(e, u), \quad t' := \sum_{u \in K} t'_u \text{ et } t_u := t'_u / t'$$

t_u est alors la probabilité, pour une réalisation du hasard, que cette réalisation se termine par l'état absorbant u .

5°) La technique proposée au 3°) admet évidemment de nombreuses autres variantes. Notamment, au lieu de considérer le couple (E', a') comme ci-dessus, on peut considérer le couple (E, a'') où a'' est défini de la façon suivante :

$$a''(u, v) := a(u, v) \text{ si } u \text{ n'appartient pas à } K$$

$$a''(u, v) := x_u \delta_{v, w} \text{ si } u \text{ appartient à } K$$

les constantes strictement positives $(x_u)_{u \in K}$ peuvent, ou non, être toutes choisies égales à 1.

La preuve de la proposition A.4 comportait une modification analogue.

D.3 Fiabilité et sureté de fonctionnement

1°) Le rôle des points absorbants est particulièrement crucial dans toutes les études de fiabilité ou de sureté de fonctionnement. Les points absorbants isolés sont les états de panne fatale pour lesquels l'évolution est complètement bloquée. Les états de dégradation irréversible correspondent

à des sous-ensembles absorbants.

Etudier la fiabilité d'un système est donc, d'un certain point de vue, calculer "les probabilités stationnaires" des états absorbants (au sens donné précédemment).

Toutefois, s'il y a un grand nombre de tels états avec des degrés divers de dégradation, la famille des probabilités stationnaires associées n'est pas immédiatement utilisable. On utilise alors de plus en plus souvent la technique de la pénalisation (cf. D.4).

2°) On considère donc un ensemble E d'états et une matrice a d'évolution associée avec les hypothèses et notations habituelles (a est positive, la diagonale de a est nulle, et d est la matrice diagonale associée à a). Soit K l'ensemble des états absorbants et $E' := E \setminus K$.

A tout point absorbant u , on associe $x_u > 0$ et $(r_{u,v})_{v \in E'}$ une famille de réels positifs telle que $\sum_{v \in E'} r_{u,v} = 1$. On pose (cf. D.2, 5°) ci-dessus) :

$$a''(u,v) := a(u,v) \text{ si } u \text{ n'appartient pas à } K$$

$$a''(u,v) := x_u r_{u,v} \text{ si } u \text{ appartient à } K.$$

En général la matrice a'' est irréductible. Dans ce cas, soit q'' la probabilité stationnaire associée à a'' .

3°) Plus x_u est faible, plus q''_u est grand. De plus, si on modifie x_u , le rapport $q''_e / q''_{e'}$ pour $e \neq u$ et $e' \neq u$ ne change pas. On est donc amené à choisir x_u d'autant plus faible que la panne associée à l'état u est grave.

Une interprétation possible est de dire que plus la panne est grave, plus elle est longue à réparer (ce qui n'est pas toujours matériellement exact).

4°) La famille $r_{u,v}$ (pour u élément de K et v élément de $E' = E \setminus K$) a elle aussi une interprétation expérimentale très précise : quelle que soit la panne, il faut bien la réparer même si la réparation doit consister à changer tout le système, c'est à dire à revenir à l'état initial où le système était neuf.

D.4 Pénalisation

1°) Une fois que l'on a "supprimé" les points absorbants (comme expliqué ci-dessus), on est ramené à la situation usuelle (cf. 2.D.1). On considère donc une matrice positive a irréductible dont tous les termes de la diagonale sont nuls. Soit d la matrice diagonale telle que $d s = a s$ où s est la

matrice uni-colonne dont tous les termes sont égaux à 1. Soit E l'ensemble des états.

2°) **A chaque état u on associe un taux de pénalisation y_u (avec $y_u \neq 0$).**

Si on ne tient pas compte des "bonifications", on choisit $y_u > 0$: y_u est d'autant plus élevé que l'état u associé est "pénalisant" c'est à dire que la situation matérielle associée est dégradée.

La pénalisation moyenne d'un sous-ensemble F de l'ensemble des états (on peut prendre $F = E$) sera alors :

$$\sum_{c \in F} q_c y_c$$

3°) Soit y la matrice diagonale (indexée par $(E \times E)$) définie par

$y_{u,v} = y_u \delta_{u,v}$. On pose $q' = q y$, $a' := y^{-1} a$ et $d' = y^{-1} d$. On a $q' a' = q' d'$ et $a' s = d' s$.

Pour le calcul de q' , modulo la constante de normalisation, tout se passe "comme si" on calculait la probabilité stationnaire du système initial sauf que l'on diminue le taux de quitter les états proportionnellement à leur pénalisation.

En général, **techniquement, il est préférable d'étudier le système $q' a' = q' d'$** plutôt que le système $q a = q d$ car il donne plus de poids à des états dont la probabilité stationnaire est faible mais le rôle crucial car associé à un état très dégradé (cf. la remarque 2.D.4, 4°).

4°) Dans ce qui précède, on n'a parlé que de pénalisation. Dans le 3°) qui précède, on peut choisir y_u positif ou négatif (mais il faut $y_u \neq 0$ quel que soit l'état u). Il est alors plus logique de choisir y_u positif et grand si l'état est excellent et y_u négatif, avec $|y_u|$ grand, si l'état est fortement dégradé.

E Equations aux dérivées partielles

E.1 Introduction

Pour alléger la présentation, nous nous limiterons au cas où on étudie des fonctions réelles $u(x,y)$ de deux variables réelles x et y : évidemment, tout ce qui suit s'étend au cas où l'on considère les fonctions de n variables, $n > 2$. On notera :

$$u'_x := \frac{\partial u}{\partial x}, \quad u''_{x,y} := \frac{\partial^2 u}{\partial x \partial y} \quad \text{etc...}$$

ce qu'on appelle les dérivées partielles première et seconde de u par rapport à x : ce paragraphe E est réservé au lecteur qui est familiarisé avec ce concept de dérivée partielle et qui a quelques connaissances élémentaires relatives aux équations aux dérivées partielles.

Soit g une fonction du couple (x,y) et des dérivées partielles de u (par exemple $g(u) = u''_{x,x} + u''_{y,y}$ ce qu'on appelle le laplacien de u). Soit h une fonction de u , x et y .

Etudier une équation aux dérivées partielles, c'est rechercher les (ou la) fonctions u satisfaisant, sur un certain domaine D , à l'équation $g(u) = h(u)$ et à certaines "conditions aux limites" : notons que le type des solutions u et la méthode à utiliser pour les déterminer dépend souvent au moins autant des conditions aux limites que des fonctions g et h .

En un certain sens, presque tous les problèmes conduisent à une équation aux dérivées partielles : leur étude est et restera donc pour des générations un domaine crucial en mathématique.

Or, il y a une **relation très étroite entre l'étude des équations aux dérivées partielles linéaires et le calcul des "probabilités stationnaires" ou, plus généralement, l'étude des systèmes en équilibre.**

Cette relation, qui est plus qu'une simple analogie, dépasse de beaucoup le petit aperçu qui va en être donné dans ce paragraphe.

Nous verrons au chapitre 9 comment, dans certains cas, le calcul des probabilités stationnaires se ramène à la résolution d'une équation aux dérivées partielles linéaire. Certes, pour qu'il en soit ainsi, il faut des hypothèses très restrictives ; cependant, quand ces hypothèses restrictives ne sont pas satisfaites, on conçoit qu'une partie des techniques - classiques en théorie des équations aux dérivées partielles - restent utilisables.

De même, la probabilité associée à un processus de diffusion, en régime stationnaire ou transitoire, satisfait à une équation aux dérivées partielles linéaires.

Dans la suite de ce paragraphe, nous allons adopter la démarche inverse, c'est à dire que nous allons montrer que la **discrétisation d'une équation aux dérivées partielles linéaires peut conduire à la résolution d'un système linéaire de la forme $q(d-a) = r$ la matrice $(d-a)$ ayant la somme de ses colonnes identiquement nulle, ce qui est le problème central de ce cours** (cf. notamment, le paragraphe 7.C). Nous nous limiterons au cas des équations aux dérivées partielles linéaires du second ordre qui ne font pas intervenir $u''_{x,y}$.

E.2 Les différences finies

Soit $u(x,y)$ une fonction réelle du couple (x,y) des variables (indépendantes) réelles x et y . Si cette fonction satisfait aux conditions adéquates de différentiabilité par rapport à la variable x on a :

à l'ordre 1 (formule des accroissements finis)

$$(6E1) \quad u(x+s,y) = u(x,y) + s u'_x + r$$

avec $\lim_{s \downarrow 0} (r/s) = 0$

à l'ordre 2 (formule de Taylor)

$$(6E2) \quad u(x+s,y) = u(x,y) + s u'_x + \frac{1}{2} s^2 u''_{x,x} + r$$

avec $\lim_{s \downarrow 0} (r/s^2) = 0$

Attention : r symbolise une fonction qui est "négligeable" mais qui dépend de l'équation considérée.

Posons pour $s \neq 0$, $t \neq 0$, s et t fixés :

$$v(j,k) := u(j s, k t)$$

Pour $x = j s$ et $y = k t$, la relation 6E1 s'écrit :

$$(6E3) \quad v(j+1,k) - v(j,k) = s u'_x + r$$

De même (toujours pour $x = j s$ et $y = k t$), la relation 6E2 implique

$$v(j+1,k) - v(j,k) = s u'_x(x,y) + \frac{1}{2} s^2 u''_{x,x}(x,y) + r_1$$

et

$$v(j-1,k) - v(j,k) = -s u'_x(x,y) + \frac{1}{2} s^2 u''_{x,x}(x,y) + r_2$$

soit, en additionnant :

$$(6E4) \quad s^2 u''_{x,x} = v(j+1,k) + v(j-1,k) - 2v(j,k) + r$$

avec $\lim_{s \downarrow 0} (r/s^2) = 0$

On a évidemment des relations analogues en intervertissant $(x, s, j, u'_x, u''_{x,x})$ et $(y, t, k, u'_y, u''_{y,y})$.

On dit que l'on a approché u'_x et $u''_{x,x}$ par des "différences finies".

E.3 Equation de type parabolique

Considérons d'abord le cas d'une équation de la forme

$$(6E5) \quad b_1(x,y) u'_x = c_1(x,y) u''_{y,y}$$

On "quadrille" le domaine D, le "pas" du quadrillage étant s dans une direction et t dans l'autre. Ceci revient à considérer les diverses valeurs du couple (x,y) pour $x = j s$ et $y = k t$, le couple (j,k) parcourant une partie de \mathbb{Z}^2 . On pose $b(j,k) := b_1(j s, k t)$ et $c(j,k) := c_1(j s, k t)$.

On remplace alors, dans 6E5, u'_x et $u''_{y,y}$ par les approximations (les différences finies) obtenues au paragraphe E.2 ci-dessus, ce qui donne

$$\begin{aligned} & [v(j,k+1) + v(j,k-1) - 2v(j,k)] c(j,k)/t^2 \\ & = [v(j+1,k) - v(j,k)] b(j,k) / s + r \end{aligned}$$

le "reste" r étant tel que r tend vers zéro quand s et t tendent vers zéro.

Si les fonctions b_1 et c_1 sont suffisamment continues, $b(j,k)$ est peu différent de $b(j+1,k)$, etc... et la relation ci-dessus peut s'écrire :

$$(6E6) \quad \begin{aligned} & [(v c)(j,k+1) + (v c)(j,k-1) - 2(v c)(j,k)] / t^2 \\ & - [(v b)(j+1,k) - (v b)(j,k)] / s = r \end{aligned}$$

le reste r (qui n'est pas nécessairement le même que dans la relation précédente) étant tel que r tend vers zéro quand s et t tendent vers zéro, s et t^2 étant du même ordre.

Pour tout couple (j,k) tel que (j s, k t) soit dans le domaine D et suffisamment loin de la frontière, on pose :

$$\begin{aligned} a[(j,k+1), (j,k)] &:= c(j,k+1) / t^2 \\ a[(j,k-1), (j,k)] &:= c(j,k-1) / t^2 \\ a[(j+1,k), (j,k)] &:= -b(j+1,k) / s \end{aligned}$$

Dans tous les autres cas , sauf près de la frontière, on pose :

$$a[(j,k), (j',k')] := 0$$

On pose, pour $(j',k') \neq (j,k)$,

$$d[(j,k), (j',k')] := 0$$

Enfin, pour $(j',k') = (j,k)$, on pose :

$$d[(j,k),(j,k)] := 2 c(j,k) / t^2 - b(j,k) / s$$

Soit v la matrice uniligne indexée par une partie de \mathbb{Z}^2 et définie par

$$v_{1,(j,k)} := v(j,k)$$

En restriction aux couples (j,k) "suffisamment à l'intérieur" du domaine associé à D , la famille de relations 6E6 s'écrit :

$$(6E7) \quad v(a-d) = r$$

où r est une matrice uniligne qui tend vers zéro quand s et t tendent vers zéro.

Evidemment, la somme des colonnes de la matrice $(a-d)$ est nulle.

E.4 Equation de type elliptique ou hyperbolique

L'étude donnée dans ce paragraphe E.4 est tout à fait analogue à celle du paragraphe E.3 qui précède. On conseille au lecteur novice de retrouver par lui-même ce qui suit à titre d'exercice (en partant de la relation 6E8 ci-dessous).

On considère le cas où l'équation est de la forme

$$(6E8) \quad b_1(x,y) u''_{x,x} = c_1(x,y) u''_{y,y}$$

On quadrille le domaine D comme en E.3 mais, alors qu'en E.3 on choisit s et t^2 du même ordre, ici on choisit s et t du même ordre. On définit b et c comme en E.3. Dans 6E8, on remplace $u''_{x,x}$ et $u''_{y,y}$ par les différences finies comme indiqué au paragraphe E.2 ci-dessus, ce qui donne :

$$\begin{aligned} & [v(j+1,k) + v(j-1,k) - 2 v(j,k)] b(j,k) / s^2 \\ & = [v(j,k+1) + v(j,k-1) - 2 v(j,k)] c(j,k) / t^2 - r \end{aligned}$$

On suppose que les fonctions b_1 et c_1 sont "suffisamment" continues ce qui permet de remplacer $b(j,k)$ par $b(j+1,k)$, etc..., le reste r étant modifié en conséquence. On obtient :

$$(6E9) \quad \begin{aligned} & - [(v b)(j+1,k) + (v b)(j-1,k) - 2(v b)(j,k)] / s^2 \\ & + [(v c)(j,k+1) + (v c)(j,k-1) - 2(v c)(j,k)] / t^2 = r \end{aligned}$$

le reste r tendant vers zéro quand s et t tendent vers zéro.

Pour tout couple (j,k) tel que (j,s,k,t) soit dans le domaine D et

suffisamment loin de la frontière, on pose :

$$a[(j+1,k), (j,k)] := -b(j+1,k) / s^2$$

$$a[(j-1,k), (j,k)] := -b(j-1,k) / s^2$$

$$a[(j,k+1), (j,k)] := +c(j,k+1) / t^2$$

$$a[(j,k-1), (j,k)] := +c(j,k-1) / t^2$$

Dans tous les autres cas , sauf près de la frontière, on pose :

$$a[(j,k),(j',k')] := 0$$

On pose, pour $(j',k') \neq (j,k)$,

$$d[(j,k),(j',k')] := 0$$

Enfin, pour $(j',k') = (j,k)$, on pose :

$$d[(j,k),(j,k)] := 2c(j,k) / t^2 - 2b(j,k) / s^2$$

On définit v comme en E.2 et en restriction aux couples (j,k) "suffisamment à l'intérieur du domaine associé à D ", la famille de relations 6E8 s'écrit :

$$(6E10) \quad v(a-d) = r$$

où r est une matrice uniligne qui tend vers zéro quand s et t tendent vers zéro. Evidemment, la somme des colonnes de la matrice $(a-d)$ est nulle.

E.5 Les conditions aux limites

Dans tout ce qui précède, nous avons considéré les équations associées à des points (j, s, k, t) dans le domaine D et "suffisamment loin de la frontière". Alors qu'il n'y a qu'un petit nombre de types d'équations aux dérivées partielles actuellement étudiées, il y a, au contraire, une très grande variété de formes possibles pour le domaine D (et donc pour la frontière) et donc une grande variété de conditions aux limites.

Toutefois, il est "facile" en théorie de discrétiser ces conditions aux bords comme on a "discrétisé l'équation centrale". Le problème est de choisir la bonne technique de discrétisation en fonction des conditions aux limites.

E.6 Les analogies

L'étude élémentaire qui précède suffit pour constater que la résolution des équations aux dérivées partielles linéaires d'ordre 2 (qui ne

font pas intervenir $u''_{x,y}$) basée sur la méthode des "différences finies" peut conduire - si on choisit une discrétisation adéquate - à la résolution d'un système linéaire $q(a-d) = r$ avec r peu différent de 0, la somme des colonnes de la matrice $(a-d)$ étant nulle.

Pour $r = 0$, on est donc exactement dans la situation extensivement étudiée dans ce cours (cf. 5.A.2). En général la matrice a est irréductible et ce point est immédiat à vérifier. De plus, **la "structure" des équations obtenues en E.3 et E.4 ci-dessus est exactement la même que celle qui apparaît dans certains problèmes de files d'attente.**

Plus précisément, considérons un réseau de files d'attente constitué de trois stations S_1, S_2, S_3 , le nombre de clients dans S_3 étant "infini" et toutes les lois de "service" étant exponentielles (c'est à dire un réseau élémentaire). Soit (j,k) l'état associé au cas où il y a j clients en S_1 et k clients en S_2 ; soit E l'ensemble des couples (j,k) "admissibles" (c'est à dire les états dont la probabilité stationnaire n'est pas nulle). On suppose que E est fini, que la "matrice d'évolution" a est irréductible et que les clients se déplacent un par un.

Le système des équations d'équilibre en régime stationnaire (cf. 2.D.2) est de la forme $q(a-d) = 0$, la matrice a étant tridiagonale. On est donc, formellement, exactement dans la même situation que pour les systèmes 6E6 et 6E9 quand $r = 0$.

Attention toutefois : l'approximation de u' et u'' proposée en E.2 n'est pas la seule possible : il y a des approximations qui ne conduisent pas à des systèmes ayant la même structure que ceux qui apparaissent dans l'étude des systèmes markoviens en régime stationnaire. Notamment on n'a pas nécessairement la somme des colonnes de $(d-a)$ égale à zéro.

De plus, dans le cas des systèmes markoviens, **la matrice a est positive**. Pour le système 6E9, c'est le cas où les fonctions b_1 et c_1 (cf. 6E8) sont de signes opposés, ce qu'on appelle le **cas elliptique** : l'exemple de base est celui de l'équation de Laplace (en dimension 2) soit $u''_{x,x} + u''_{y,y} = 0$. Pour le système 6E6, la matrice a est positive si les fonctions b_1 et c_1 sont de signes opposés (cf. 6E5) ; si ce n'est pas le cas, on change x en $-x$ (ce qui revient à approcher u'_x par $[v(j,k) - v(j-1,k)] / s$ au lieu d'utiliser 6E3). donc, dans tous les cas, si le produit $b_1 c_1$ ne change pas de signe et si l'équation est de **type parabolique** (en dimension 2) on peut se ramener au cas où **a est positive**. Notons que, dans ce cas, la structure de la matrice a est particulièrement simple (cf. 5.E.5, 2°).

E.7 Les différences

Ces analogies ne doivent pas occulter de profondes différences. Dans le cas des systèmes markoviens, il y a une très grande variété de structures possibles pour les matrices a "au niveau de l'équation centrale" (comme en 6E6 et 6E9), le nombre de paramètres tels que j et k ci-dessus est grand. Par contre, les conditions aux bords sont souvent assez simples.

Au contraire, dans le cas des équations aux dérivées partielles, le nombre "d'inconnues" est peu élevé, la structure de l'"équation centrale" est donc relativement simple. Par contre, il y a une très grande variété de conditions aux bords.

Enfin, et surtout, la résolution des systèmes linéaires tels que 6E6 ou 6E9 n'est qu'une étape de l'étude ; il faut aussi étudier la limite de la solution quand s et t tendent vers zéro.

E.8 Cas avec second membre

L'essentiel de ce cours porte sur le cas "sans second membre", c'est à dire l'étude du système $q(a-d) = 0$: la raison essentielle en est que, techniquement, dans le cas des systèmes markoviens, on ne sait "presque rien" faire dans le cas "avec second membre", c'est à dire le cas d'un système de la forme $q(a-d) = g$ où g est "connue" et q "inconnue" (cf. toutefois, 5.E.6 ci-dessus, ce chapitre, le paragraphe 7.C et le chapitre 9).

Pour autant, le cas "avec second membre" est crucial pour l'étude des systèmes markoviens : régime transitoire, stabilité, points absorbants, calculs intermédiaires, etc...

De même, dans le cas des équations aux dérivées partielles linéaires de la forme

$$b_1 u'_x - c_1 u''_{y,y} = f \quad \text{ou} \quad b_1 u''_{x,x} - c_1 u''_{y,y} = f$$

la discrétisation à l'aide des différences finies (comme expliqué ci-dessus) conduit à la résolution d'un système linéaire de la forme $q(a-d) = g + r$ où r est "négligeable" par rapport au "second membre" g : ce second membre est "petit" mais non négligeable.

F. Matrices de Léontief

F.1 Introduction

En 1941, Wassily W. Léontief a publié "The structure of american economy, 1919-1929" (cf. [Leo]). L'objet de cet ouvrage est de préciser les

relations qui existent entre diverses quantités économiques. Ces relations sont essentiellement linéaires et peuvent donc être interprétées en termes de matrices.

Le but de ce paragraphe n'est pas de résumer la partie mathématique du livre de Léontief ; il se propose simplement d'attirer l'attention du lecteur sur le fait que les équilibres économiques peuvent, en première approximation, s'interpréter en termes de solution p d'un système linéaire $p(a-d) = q$ (éventuellement avec $q = 0$) où les matrices a et d sont positives, la matrice d étant diagonale, ce qui est le problème central étudié dans ce cours. Ceci n'a rien de surprenant puisqu'il doit y avoir équilibre entre les "flux d'entrée" et les "flux de sortie". En général on se ramène au cas où la matrice d est la matrice unité.

Evidemment, ceci n'est qu'une toute petite partie de l'utilisation des mathématiques en économie. Notamment, une étude fine de la solution à l'équilibre utilise la notion d'équilibre optimal ce qui fait appel - aussi bien du point de vue mathématique que du point de vue économique - à des théories beaucoup plus sophistiquées que les modèles de Léontief.

F.2 Equilibre des prix

On considère divers secteurs ou divers types de production économiques. Soit I l'ensemble de ces secteurs ou de ces types de production. La répartition choisie peut être plus ou moins fine : par exemple, un élément i de I peut être associé à la production de blé dans une région ou à l'ensemble de la production agricole d'un pays.

Pour chaque élément i de I , on choisit une unité de production, par exemple le millier de quintaux de blé. S'il s'agit d'une production diversifiée, l'unité est souvent exprimée en termes monétaires : par exemple, l'unité de référence peut être la production agricole équivalente à un million de dollars au premier janvier 1989.

Si on se propose d'étudier l'équilibre des prix, pour tout élément i de I on appelle p_{i0} le prix de l'unité de production : évidemment, dans l'exemple ci-dessus, si on se propose d'étudier l'évolution du prix au cours de l'année 1989, le prix p_a de l'équivalent agricole d'un million de dollars au premier janvier 1989 ne va pas rester égal à 1 durant l'année.

Dans ce cas, déterminer la matrice de Léontief, c'est déterminer, pour tout couple (i,j) d'éléments de I , le nombre (réel) a_{ij} d'unités de la production i nécessaires pour produire une unité de la production j .

Puisqu'il y a équilibre, on doit avoir :

quel que soit i élément de I :

$$p_i = \sum_j p_j a_{j,i}$$

Ce système d'équations linéaires peut aussi s'écrire $p(d-a) = 0$ où p est la matrice uniligne $p := (p_i)_{i \in I}$, d est la matrice unité $(I \times I)$ et a est considérée comme une matrice carrée $(I \times I)$.

Notons que la matrice a n'est pas nécessairement triangulaire : par exemple si i est l'indice associé au fuel et j l'indice associé au transport, on a à la fois $a_{i,j} > 0$ et $a_{j,i} > 0$.

De plus - et ceci est sous-jacent mais non explicité dans [Leo] - considérons une décomposition qui permet d'étudier les salaires et les revenus des capitaux séparément des autres types de production. Soit s et c les indices associés respectivement aux salaires et aux capitaux.

Les coefficients $a_{i,s}$ et $a_{i,c}$ pour $i \neq s$ et $i \neq c$ ne sont alors pas du tout de la même nature économique que les autres coefficients. Par exemple $a_{i,s}$ est le "nombre" d'unité de production i qui est acheté avec le pourcentage **moyen** - pourcentage moyen affecté à la production i par l'ensemble des salariés - de l'unité de salaire.

Evidemment, la matrice a doit être telle que le système $p(a-d) = 0$ ait une solution non nulle, c'est à dire que le déterminant de $(a-d)$ est égal à 0 : ceci correspond à la "cohérence économique" des coefficients $a_{i,j}$.

F.3 Equilibre des flux économiques

Supposons qu'au lieu d'étudier l'équilibre des prix, on s'intéresse à **l'équilibre des flux économiques**. Pour chaque élément i de I , soit q'_i l'équivalent monétaire du "flux de sortie" de i , c'est à dire la production - exprimée en millions de dollars par exemple - du secteur économique i . Pour tout couple (i,j) d'éléments de I , soit $a'_{i,j}$ le pourcentage de la production du secteur i qui est utilisée par le secteur j . La matrice a' de terme général $a'_{i,j}$ est une matrice stochastique (cf. 1.C.3). On a évidemment (si on a pris en compte tous les secteurs) :

$$q'_i = \sum_j q'_j a'_{j,i}$$

ce qui est encore de la forme $q'(d'-a') = 0$ où d' est la matrice unité.

Le livre de Léontief étant écrit dans un langage plus économique que

mathématique, il y a beaucoup d'autres façons de présenter les "matrices de Léontief" : cf., notamment le paragraphe 2.1 de [Sen] où on considère un système $p(a-d) = q$ avec $q \neq 0$.

F.4 Aspect descriptif

Le premier intérêt des matrices de Léontief est descriptif : elles permettent de mieux cerner les relations entre les prix, les flux économiques, etc... Cet aspect suscite peu de commentaires mathématiques sauf que les matrices a et d ou a' et d' satisfont à des propriétés de cohérence analogues à celles que l'on a rencontrées dans l'étude des systèmes markoviens. Il faut aussi noter que, contrairement aux autres exemples étudiés dans ce cours, la matrice a (ou a') n'est pas, en général, tridiagonale par blocs.

F.5 Sensibilité

Léontief a bien noté que la modélisation qu'il propose n'a pas qu'un intérêt descriptif : elle permet aussi d'étudier relativement simplement l'évolution des paramètres les uns par rapport aux autres. Par exemple elle permet d'étudier l'influence qu'un gain de productivité - c'est à dire la diminution d'un coefficient a_{ij} avec $j \neq s$ et $j \neq c$ - peut avoir sur l'augmentation des pouvoirs d'achat et/ou de la rémunération du capital en fonction de la répartition du gain de productivité entre les trois secteurs : salaires, dividendes et investissements.

D'ailleurs, il n'est peut-être pas interdit de penser qu'un jour les revendications chiffrées syndicales porteront essentiellement non plus sur le salaire mais sur cette répartition entre l'augmentation du pouvoir d'achat et la rémunération du capital.

Or, tant que les coefficients a_{ij} varient peu, l'approximation linéaire - qui est à la base de la modélisation de Léontief - est très satisfaisante. L'étude de la sensibilité, c'est à dire l'évolution relative des divers paramètres, peut alors être effectuée à partir de la théorie générale des déterminants.

En fait, les matrices considérées ont une structure très particulière ce qui permet d'effectuer une étude spécifique très précise de la sensibilité : ce point est au coeur du chapitre 7 qui, rappelons-le, est le chapitre central de ce cours.

G. Conclusion

Comme on l'a noté dans l'introduction, le point essentiel à retenir de ce chapitre 6 se situe au niveau de la méthodologie.

D'une part, des situations apparemment sans relation, ou n'ayant qu'une relation éloignée, avec l'étude d'un régime stationnaire, se ramènent souvent à la situation introduite en 2.D.1. Ceci est vrai sur le plan expérimental (suppression des états absorbants, pénalisation, E.D.P., etc....). C'est également vrai sur le plan mathématique : l'étape essentielle de la résolution d'un système linéaire très général peut être la résolution d'un système comme ceux introduits en 2.D.1 (il suffit de choisir un "bon" splitting).

D'autre part, dans l'étude des processus de Markov, les "opérations chirurgicales" constituent une technique classique : mort d'un processus (point absorbant), recollement ou décomposition de processus (comme dans la chaîne incluse), changement dans l'échelle des temps (étude de $y^{-1}a$ au lieu de a), etc... Quand on étudie un processus en régime stationnaire, ces "opérations" sont en général associées à des relations matricielles "simples" mais que l'interprétation probabiliste peut aider à découvrir. En général, le champ d'application de ces propriétés matricielles dépasse très largement le cadre probabiliste.

Chapitre 7

Graphes

A. Réseaux fluviaux

A.1 Graphe et matrice d'évolution

1°) Un ensemble (fini) E étant donné, soit a une fonction définie sur $(E \times E)$; l'essentiel de ce cours porte sur l'étude et la résolution d'un système linéaire associé à la "matrice" a . Plus précisément, on suppose que, pour tout élément u de E , $a(u,u) = 0$ et on pose

$$d(u,v) := \delta_{u,v} \sum_{w \in E} a(u,w)$$

On étudie alors le système $q(a-d) = 0$ (cas "sans second membre" : cf. 2.D.1, 5.A.2, etc...)

et le système $y(a-d) = x$ (cas "avec second membre" : cf. 5.E.6, 6.A.2, etc...)

2°) En fait, les techniques et les exemples proposés dans ce cours privilégient le cas où la matrice a est positive : notamment, dans ce cas, la solution normalisée q du système $q(a-d) = 0$ peut être la probabilité stationnaire d'un processus markovien homogène dont l'évolution est régie par la matrice a (cf. 2.C.4).

3°) Un intermédiaire "technique", souvent indispensable, pour étudier les systèmes linéaires évoqués ci-dessus est la "théorie des graphes" (cf. [Ber]). En fait, cette "théorie" intervient de plus en plus dans presque tous les domaines, notamment en mathématiques et en informatique. Ce qui suit ne fait appel qu'à une petite partie de cette théorie.

4°) Le "graphe" associé à la matrice a admet E comme ensemble de points. L'ensemble des arcs ordonnés est l'ensemble des couples (u,v) d'éléments de E tels que $a(u,v) \neq 0$.

Evidemment, si on a à la fois $a(u,v) \neq 0$ et $a(v,u) \neq 0$, le graphe comprend l'arc (u,v) et l'arc (v,u) .

5°) On a noté (cf. 2.D.3) l'importance du cas où la matrice a est irréductible : cette propriété a une interprétation très simple en termes de graphe. En effet, elle revient à dire que si on parcourt le graphe suivant ses arcs (en respectant leur orientation) on peut, à partir de n'importe quel point u atteindre n'importe quel autre point v .

Avec cette "visualisation" il est "facile" d'étudier la décomposition des matrices a qui ne sont pas irréductibles. Plus précisément, les sous-ensembles "absorbants" sont les sous-ensembles d'états "équivalents" au sens suivant : u et v sont équivalents si on peut aller de u à v et de v à u en parcourant le graphe comme indiqué plus haut (un sous-ensemble absorbant peut être réduit à un point). A chaque sous-ensemble absorbant correspond une sous-matrice irréductible.

Nous ne détaillerons pas davantage cette partie de l'étude que l'on peut trouver dans de très nombreux ouvrages et dont la résolution est presque toujours quasi-triviale quand on considère un exemple issu d'une modélisation de la réalité expérimentale.

6°) Une deuxième utilisation de la "technique" des graphes, quand on cherche la solution q du système $q(a-d) = 0$, est tout simplement de **dessiner**, en tout ou en partie, le **graphe associé à a** . Si le problème étudié est très simple, il peut être intéressant de dessiner le graphe complet.

En général, on ne dessine que certains sous-graphes ; par exemple, en 5.C.2, pour les deux exemples évoqués (files PH/M/1 et file avec buffer), l'ensemble des états est naturellement plongé dans \mathbb{N}^2 ; le graphe associé est donc "facile" à "dessiner". Pour chaque type d'état u , u fixé ; il est commode de dessiner le sous-graphe associé à tous les états v tels que $a(v,u) \neq 0$. La figure 7.a représente ce dessin dans le cas de la file avec buffer pour un état $u = (j,k)$ tel que $0 < j < J$.

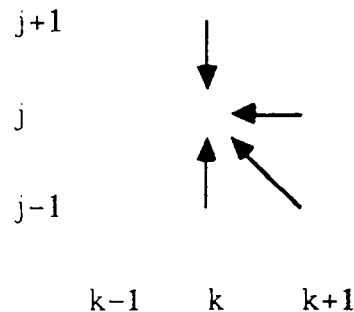


Figure 7.a

Quand la situation est complexe (par exemple, cas de [Sag]), il est opportun d'attribuer une couleur à chaque type de transition : le tracé du graphe constitue alors un auxiliaire technique presque indispensable.

7°) Dans la suite de ce chapitre, nous allons utiliser une partie de la théorie des graphes de façon beaucoup plus profonde et donc plus difficile à assimiler.

Attention donc : ce chapitre est d'un haut niveau théorique. Le lecteur plus attiré par les applications que par la théorie sous-jacente peut,

dans un premier temps, se contenter de comprendre les énoncés des théorèmes de base.

Toutefois, ce chapitre est, à tous points de vue, au coeur du cours ; comme c'est souvent le cas quand il s'agit d'une étude théorique, sa compréhension "en profondeur" n'est pas indispensable au niveau des applications : par contre, c'est la "clef de voûte" qui permet de mieux voir comment s'ordonnent les autres chapitres.

Avant de donner la définition d'un réseau fluvial, il n'est peut-être pas inutile de préciser un point de vocabulaire.

A.2 Vocabulaire et définition

La notion de "**définition mathématique**", contrairement à ce qu'on pourrait croire, ne fait pas, actuellement l'objet d'un consensus précis et unanime auprès des mathématiciens. Le consensus, pas toujours respecté d'ailleurs, est constitué des limites entre l'acceptable et l'inacceptable. Prenons quelques exemples :

Nous avons, précédemment, dit qu'une matrice réelle $(a_{i,j})$ est une famille de nombres réels indexée par $(E \times E)$, ou encore une fonction réelle définie sur $(E \times E)$: cette définition est acceptable en ce sens que la connaissance de la-dite fonction suffit pour caractériser tout ce qui est "attaché" à une ou plusieurs matrices (déterminant, produit de matrice, etc...).

Certains auteurs préfèrent dire qu'une matrice est "le représentant d'une application linéaire pour la relation d'équivalence..." ; est-ce plus clair ? Est-ce convaincant quand on étudie la notation par blocs ou, pire, la forme bilinéaire associée à une matrice symétrique ?

Le cas des séries est encore plus angoissant : une série $(u_n)_{n \geq 0}$ à termes réels est caractérisée par l'application f définie sur \mathbb{N} par $f(n) := u_n$; c'est donc une suite ; mais la série de terme général u_n est dite convergente

si la suite de terme général $s_n := \sum_{k=0}^n u_k$ est convergente. L'artifice utilisé dans ce cas par le prestigieux Bourbaki ne change rien au fond du problème.

Enfin le cas des variables aléatoires ou, a fortiori, des processus relève du casse-tête chinois. On peut toutefois noter que confondre la notion de variable aléatoire et celle d'application mesurable est une erreur de logique inacceptable si on utilise des expressions telles que l'espérance d'une variable aléatoire : en effet, l'espérance de X n'est déterminée que si on connaît la loi de X ce qui est beaucoup plus que de savoir que X est une

application mesurable.

En fait, il faudrait distinguer entre la définition du "concept" et celle de l'"objet". Quand on utilise le mot matrice cela sous-entend que le produit de deux matrices n'est pas défini comme le produit des fonctions associées, alors qu'au niveau de l'objet, une matrice indexée par $(E \times E)$ est caractérisée par une fonction définie sur $(E \times E)$. Si les mathématiciens ne clarifient pas ce point, les informaticiens le feront (et en tireront les bénéfices !) : l'étude d'un langage tel que Lisp en montre la nécessité.

De même, dans le paragraphe A.3 qui suit, un réseau fluvial (comme une arborescence) est caractérisé, en tant qu'objet, par une fonction r . Par contre, le concept de réseau fluvial ne prend son véritable sens qu'une fois que l'on a défini la mer d'un réseau fluvial, la notion d'état antérieur, pour un réseau fluvial, la notion de réseau fluvial propre, etc... De ce point de vue, un réseau fluvial n'est pas une arborescence, même en un sens "généralisé". (cf., par exemple, [Ber]).

A.3 Réseau fluvial

Etant donné E ensemble des états, E fini, un réseau fluvial - "en tant qu'objet" - est une application r définie sur E et à valeurs dans E .

Le graphe ordonné associé à un réseau fluvial r est l'ensemble des arcs (des couples) $(u, r(u))$ pour u élément de E . On dira que $r(u)$ est le successeur de u . On dira que u est antérieur (resp. postérieur) à v s'il existe un entier $k \geq 0$ tel que $v = r^k(u)$ (resp. $u = r^k(v)$), avec la convention $r^0(u) = u$ (u est donc antérieur et postérieur à u).

La mer associée à un réseau fluvial r est l'ensemble des éléments u de E pour lesquels il existe $k > 0$ tel que $r^k(u) = u$.

On dira qu'un élément u de E est **isolé** si $u = r(u)$.

On dira qu'un réseau fluvial r est **propre** si sa mer est un ensemble de points isolés. Si r est un réseau fluvial propre, quel que soit v élément de E , $r^k(v)$ est fixe pour k assez grand ; on peut donc poser :

$$r^*(v) := \lim_{k \rightarrow \infty} r^k(v)$$

Evidemment, $r^*(v)$ est un élément de la mer (celui auquel aboutit un "fleuve" qui passe en v).

Un sous-ensemble A non vide et non réduit à un élément de E sera appelé **anneau** (pour le réseau fluvial r) si chaque élément u de A est antérieur (et donc postérieur) à chaque élément v de A . Un anneau est contenu dans la mer.

La mer d'un réseau fluvial est constituée d'anneaux et/ou de points isolés.

Si r est un réseau fluvial (sur E) et a une fonction réelle (en général positive) définie sur $(E \times E)$, on pose :

$$\pi(a, r) := \prod_{u \in Z} a(u, r(u)) \quad \text{avec}$$

$$Z := \{ u : u \in E, u \neq r(u) \}$$

On note que $\pi(a, r)$ ne dépend pas des valeurs de $a_{u, u}$.

Le cas le plus important est le cas où la mer du réseau fluvial r est réduite à un point e : dans ce cas on dira que r est un réseau fluvial de mer e (au lieu de dire de mer $\{e\}$). Les autres cas sont des intermédiaires techniques.

A.4 Cas sans second membre

Théorème : Soit E un ensemble fini comprenant au moins deux éléments et a une fonction (réelle ou complexe) définie sur $(E \times E)$ telle que, pour tout élément u de E , $a(u, u) = 0$. Pour tout élément u de E , soit $A(u)$ l'ensemble des réseaux fluviaux propres de mer u et soit $\alpha(u)$ défini par :

$$(7A1) \quad \alpha(u) := \sum_{r \in A(u)} \pi(a, r)$$

On a alors, pour tout élément u de E :

$$(7A2) \quad \alpha(u) \sum_{v \in E} a(u, v) = \sum_{v \in E} \alpha(v) a(v, u)$$

Notamment, si la matrice a est positive irréductible (cf. 2.B.3), la famille α est une solution de 4A1 : c'est donc, à une constante multiplicative près, la probabilité stationnaire associée à la matrice d'évolution a (cf. 2.D). Si a est une matrice positive, la famille α est strictement positive si et seulement si a est irréductible.

Preuve :

1°) Soit u un élément de E . Soit $F(u)$ l'ensemble des réseaux fluviaux dont la mer est constituée d'un anneau qui contient u . On pose :

$$f(u) := \sum_{r \in F(u)} \pi(a, r)$$

On se propose de démontrer que chaque terme de la relation 7.A.2

est égal à $f(u)$. Cette démonstration est typique ; pour en faciliter la compréhension, on conseille vivement au lecteur d'observer simultanément un dessin associé (figure 7b).

2°) Soit $F'(u)$ l'ensemble des couples (r', v') où r' est un réseau fluvial de mer u et v' un élément de E différent de u .

Soit s l'application, définie sur F' , qui au couple (r', v') associe la fonction r définie par $r(u) := v'$ et, pour $w \neq u$, $r(w) := r'(w)$; $r = s(r', v')$ est un élément de $F(u)$.

3°) Soit t l'application définie sur F qui à r élément de F associe le couple (r', v') défini comme suit : $r'(u) := u$ et, pour $w \neq u$, $r'(w) := r(w)$; $v' = r(u)$ (v' est sur l'anneau associé à r).

4°) Pour tout élément r de F , on a

$$s[t(r)] = r$$

De même, pour tout élément x de F' , $t[s(x)] = x$.

Donc s (et t) établit une bijection entre les ensembles $F(u)$ et $F'(u)$ ce qui implique :

$$f(u) = \alpha(u) \sum_{v \in E} a(u, v)$$

5°) Une façon de "visualiser" cette partie de la preuve est de dire que, pour chaque élément r de F , on "coupe" l'anneau associé "après u ".

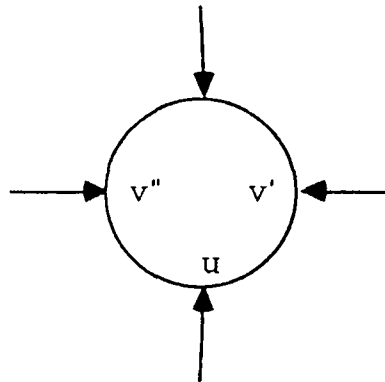


figure 7b

6°) Soit $F''(u)$ l'ensemble des couples (r'', v'') avec v'' élément de E distinct de u et r'' réseau fluvial de mer v'' . En "coupant" l'anneau associé à chaque élément de $F(u)$ "avant u ", on construit une bijection de $F(u)$ sur $F''(u)$ ce qui implique

$$f(u) = \sum_{v \in E} \alpha(v) a(v, u)$$

7°) Si la matrice a est positive et irréductible, pour tout élément u de E , il existe un réseau fluvial r de mer u tel que $\pi(a,r) > 0$ (on construit $r(G)$ par récurrence sur le nombre d'états éléments de G). La famille α est donc strictement positive.

8°) Si la matrice a est positive et si la solution α du système 7A2 est strictement positive, cela implique que, pour chaque élément u de E , il existe un réseau fluvial r de mer u tel que $\pi(a,r) > 0$. Ceci implique que, à partir de tout état v distinct de u , "on peut atteindre u ". Ceci étant vrai quel que soit u , la matrice a est irréductible.

A.5 Irrigabilité

Définition : Etant donné un ensemble fini E , une matrice positive a indexée par $(E \times E)$ et deux éléments u et v de E , on dira que v est a -irrigable par u si $u = v$ ou s'il existe une séquence $(w_j)_{1 \leq j \leq k}$ telle que $w_1 = u$, $w_k = v$ et, quel que soit j , $1 \leq j < k$, $a(w_j, w_{j+1}) > 0$. Ce qui équivaut à dire qu'il existe un entier i tel que $(a^i)_{u,v} > 0$.

Remarques

1°) Par définition (cf. 2.B.3), la matrice positive a est irréductible si et seulement si, quels que soient u et v éléments de E , v est a -irrigable par u .

2°) Soit α la fonction associée à la matrice positive a comme défini en A.4 ; on a $\alpha(u) > 0$ si et seulement si u est a -irrigable par tous les éléments v de E (ceci se prouve comme le 8°) de A.4).

3°) Si v est a -irrigable par u et w est a -irrigable par v , alors w est a -irrigable par u .

A.6 M-matrice inversible

Théorème : Soit E un ensemble fini. Soit d et c deux matrices positives indexées par $(E \times E)$, la matrice d étant diagonale et tous les termes de la diagonale de c étant nuls.

On suppose que, pour tout élément u de E , $d_{u,u} \geq \sum_{v \in E} c_{u,v}$.

Soit F l'ensemble des éléments u de E tels que $d_{u,u} > \sum_{v \in E} c_{u,v}$.

Soit G l'ensemble des éléments u de E tels qu'il existe un élément v de F (qui peut dépendre de u) qui est c -irrigable par u .

Alors la matrice $m := (d - c)$ est inversible si et seulement si $G = E$. De plus, dans ce cas $G = E$, la matrice inverse de m est positive c'est à dire que m est une M-matrice.

Remarques :

Ce théorème précise le théorème 4.41 de [GoM]. Notons aussi que la propriété $G = E$ se "visualise" très bien en termes de théorie des graphes.

Preuve : condition suffisante

1°) On suppose que $G = E$ et on se propose de prouver que l'on peut appliquer la proposition 6.A.4. Comme dans cette proposition, appelons s la matrice uni-colonne indexée par $(\{1\} \times E)$ et dont tous les termes sont égaux à 1.

On pose $f := (d-c) s$: cette matrice uni-colonne est positive et le terme f_u de la "ligne u " de f est différent de zéro si et seulement si u appartient à F .

2°) Soit w un élément fixé de E . Soit x la matrice (indexée par $(E \times E)$ et qui dépend de w) qui est égale à $(d-c)$ sauf sa "colonne w " qui est telle que $x s = 0$: c'est aussi la matrice notée $(d-c-\Phi(w,f))$ dans la proposition 6.A.4.

Soit a la matrice, indexée par $(E \times E)$, dont la diagonale est nulle et qui est définie par $a_{u,v} := -x_{u,v}$ pour $u \neq v$ (cette matrice dépend de w). Cette matrice a est positive. Si u appartient à F avec $u \neq w$, $a_{u,w} > 0$: w est donc a -irrigable par u .

3°) De plus, si i et j sont deux éléments distincts de E et si j est c -irrigable par i , j est aussi a -irrigable par i puisque $a_{h,k} \geq c_{h,k}$ pour $h \neq k$.

La propriété $E = G$ implique donc que, quel que soit u élément de E , il existe v élément de F qui est a -irrigable par u .

4°) Considérons d'abord le cas où w appartient à F et soit v un élément quelconque de E ; il existe v' élément de F qui est a -irrigable par v (cf. la fin du 3°) et w est a -irrigable par v' (cf. la fin du 2°) donc w est a -irrigable par v : ceci étant vrai quel que soit v , $\alpha(w) > 0$. Puisque w appartient à F , si on considère α comme une matrice uni-ligne, on a $\alpha f > 0$.

5°) Considérons maintenant le cas où w n'appartient pas à F ; il existe un élément u de F qui est a -irrigable par w (cf. le 3°) ; fixons u ; on prouve comme au 4°) que w est a -irrigable par n'importe quel élément v de E ; il en est donc de même pour u et on a $\alpha(u) > 0$. Puisque u appartient à F , comme au 4°) , si on considère α comme une matrice uniligne, on a $\alpha f > 0$.

6°) Les 4°) et 5°) montrent que, dans tous les cas, $\alpha f > 0$ et on peut appliquer la proposition 6.A.4 ; plus précisément, pour chaque valeur de w , α est, à une constante multiplicative près, la "ligne w " de la matrice inverse de $(d-c)$. Cette matrice inverse est donc ainsi construite et elle est positive, ce qui prouve que m est une M -matrice.

7°) Condition nécessaire

On suppose que $E' := E \setminus G$ n'est pas vide. Soit s' , d' et c' les restrictions de s , d et c à E' . On a $(d'-c')s' = 0$: la "sous-matrice" $(d'-c')$ n'est donc pas inversible. La famille des lignes de $(d'-c')$ est donc linéairement dépendante (i.e. ce n'est pas une famille libre). Or,

$(d-c)_{u,v} = 0$ si u appartient à E' et v n'appartient pas à E' : la famille des lignes de $(d-c)$ associées à E' est donc linéairement dépendante et la matrice $(d-c)$ ne peut pas être inversible.

B. Relations de base

Dans tout ce paragraphe B, on suppose que E est un ensemble fini qui a au moins trois éléments et que a est une fonction (réelle ou complexe) définie sur $(E \times E)$ telle que $a(u,u) = 0$ pour tout élément u de E .

B.1 Notation β

Soit (u,v,w) un triplet d'éléments de E . Pour $u = v$ et pour $u = w$ on pose :

$$\beta(u,u,w) := \beta(u,v,u) := 0$$

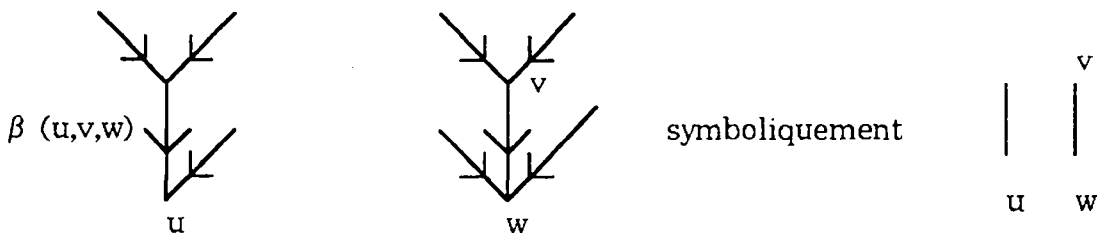
Pour $u \neq v$ et $u \neq w$, soit $B(u,v,w)$ l'ensemble des réseaux fluviaux propres r qui admettent les deux points isolés u et w comme mer et tels que v est antérieur à w . On pose :

$$(7B1) \quad \beta(u,v,w) := \sum_{r \in B(u,v,w)} \pi(a,r)$$

On vérifie, en raisonnant comme en A.4, que l'on a, pour $u \neq w$:

$$(7B2) \quad \alpha(w) := \sum_{v \in E} \beta(u,v,w) a(u,v)$$

(pour chaque réseau fluvial élément de $A(w)$, on "coupe après u ", A étant défini comme en A.4).



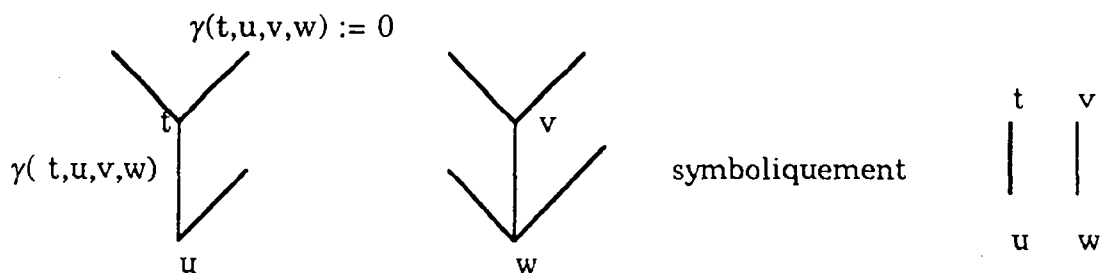
Notons que, si $B(u,v,w)$ est vide, on a posé $\beta(u,v,w) := 0$.

B2 Notation $\gamma(t,u,v,w)$

Soit (t,u,v,w) quatre éléments de E avec $t \neq v$, $t \neq w$, $u \neq v$ et $u \neq w$. Soit $C(t,u,v,w)$ l'ensemble des réseaux fluviaux propres dont la mer est constituée des deux points isolés u et w et tels que t est antérieur à u et v est antérieur à w . On pose :

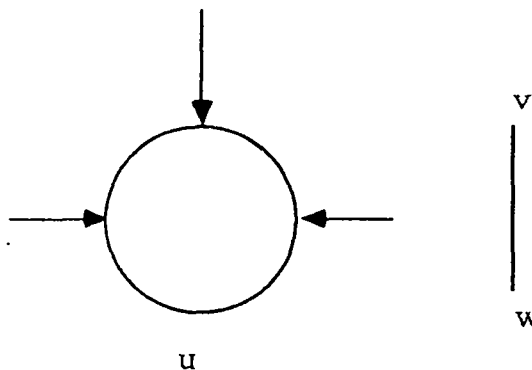
$$(7B3) \quad \gamma(t,u,v,w) := \sum_{r \in C(t,u,v,w)} \pi(a,r)$$

Si on a soit $t = v$, soit $t = w$, soit $u = v$, soit $u = w$, (c'est à dire si $C(t,u,v,w)$ est vide), on pose



Lemme : Soit u, v et w trois éléments de E . On a :

$$(7B4) \quad \sum_t \gamma(t,u,v,w) a_{u,t} = \sum_t \gamma(u,t,v,w) a_{t,u}$$



Preuve :

La preuve de ce lemme est tout à fait analogue à la preuve de A.4. Si $u = v$ ou $u = w$, la relation 7B4 est évidente. Supposons donc $u \neq v$ et $u \neq w$. Soit F l'ensemble des réseaux fluviaux r dont la mer comprend d'une part un anneau qui contient u , d'autre part l'élément isolé w , v étant antérieur à w . On pose

$$f(u,v,w) := \sum_{r \in F} \pi(a,r)$$

On prouve comme en A.4 que $f(u,v,w)$ est égal au membre de gauche de 7B4 (on coupe l'anneau après u) et au membre de droite de 7B4 (on coupe l'anneau avant u).

B.3 Notation $\gamma'(u,t,v,w)$

Soit u, t, v et w quatre éléments de E . Pour $u = t$ ou $u = v$ ou $u = w$ on pose $\gamma'(u,t,v,w) := 0$. Pour $u \neq t$ et $u \neq v$ et $u \neq w$, soit $C'(u,t,v,w)$ l'ensemble des réseaux fluviaux propres dont la mer est constituée des deux points isolés u et w et pour lesquels t et v sont antérieurs à w . On pose :

$$(7B5) \quad \gamma'(u,t,v,w) := \sum_{r \in C'(u,t,v,w)} \pi(a,r)$$

Evidemment, $\gamma'(u,v,t,w) = \gamma'(u,t,v,w)$

Lemme : Quel que soit (u,t,v,w) quadruplet d'éléments de E , on a :

$$(7B6) \quad \beta(u,v,w) = \gamma(t,u,v,w) + \gamma'(u,t,v,w)$$

$$\begin{array}{ccc} & v & \\ | & | & \\ u & w & \end{array} = \begin{array}{ccc} & t & v \\ | & | & \\ u & w & \end{array} + \begin{array}{ccc} & & tv \\ | & | & \\ u & w & \end{array}$$

$$(7B7) \quad \gamma(u,t,v,w) - \gamma(v,t,u,w) = \beta(t,v,w) - \beta(t,u,w)$$

Vérification

1°) Soit r un réseau fluvial dont la mer est constituée des éléments isolés u et w et tel que v est antérieur à w ; on a soit t antérieur à u , soit t antérieur à w , ces deux possibilités étant exclusives. Ceci implique 7B6.

2°) La relation 7B6 implique

$$\gamma(u,t,v,w) = \beta(t,v,w) - \gamma'(t,u,v,w)$$

et le premier membre de 7B7 vaut :

$$\beta(t,v,w) - \gamma'(t,v,u,w) - \gamma(v,t,u,w)$$

ce qui est égal au deuxième membre de 7B7 (on utilise à nouveau la relation 7B6).

B.4 Notation $\beta'(v,u,w)$

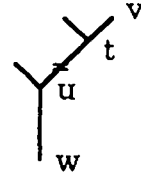
Soit u, v et w trois éléments de E . Soit $B'(v,u,w)$ l'ensemble des réseaux fluviaux propres de mer w tels que v est antérieur à u . Si $B'(v,u,w)$ est vide on pose $\beta'(v,u,w) := 0$. Sinon, on pose :

$$(7B8) \quad \beta'(v,u,w) := \sum_{r \in B'(v,u,w)} \pi(a,r)$$

Lemme : Quel que soit (u,v,w) triplet d'éléments de E , avec $u \neq v$, on a :

$$(7B9) \quad \beta'(v,u,w) = \sum_t \gamma(v,t,u,w) a_{t,u}$$

$$\begin{array}{c} v \\ | \\ u \\ | \\ w \end{array} = \sum_t \begin{array}{c} v \\ | \\ t \\ | \\ w \end{array} \begin{array}{c} u \\ | \\ t \\ | \\ w \end{array} a_{t,u}$$



$$(7B10) \quad \alpha(w) = \beta'(v,u,w) + \sum_t \gamma'(u,t,v,w) a_{u,t}$$

et

$$(7B11) \quad \alpha_u = \sum_t \beta(u,v,t) a_{t,u}$$

Vérification

1°) Soit r un élément de $B'(v,u,w)$; en coupant avant u en sorte que v ne soit plus antérieur à u (il y a une seule telle coupure possible), on obtient un couple (r',t) unique où t est un élément de E et r' un élément de $C(v,t,u,w)$. On vérifie comme en A.4 que l'on a une bijection ce qui implique 7B9.

2°) Soit r un réseau fluvial de mer w . On a, de façon exclusive, l'une des deux possibilités suivantes : ou bien v est antérieur à u ce qui donne $\beta'(v,u,w)$ quand on fait la somme sur tous les réseaux fluviaux possibles r de ce type, ou bien v n'est pas antérieur à u (ce qui est impossible si $u = w$). Dans ce deuxième cas, on coupe après u (ce qui "sépare" u et v) et on achève la preuve de 7B10 comme en A.4.

3°) La preuve effectuée au 1°) qui précède vaut même si $u = w$ ce qui donne la relation 7B11.

B.5 Théorème

Soit (u,v,w) un triplet d'éléments de E avec $u \neq v$.

On pose $\lambda(u,w) := \sum_t \beta_{t,u,w} a_{t,u}$.

On a :

$$(7B12) \quad \beta_{u,v,w} \sum_t a_{u,t} = \sum_t \beta_{t,v,w} a_{t,u} + \alpha(w) - \lambda(u,w)$$

Preuve :

Cette relation se vérifie directement quand $u = w$ (cf. 7B11 et 7B6). On suppose donc $u \neq w$.

On pose $x := \beta(u,v,w) \sum_t a_{u,t}$

$$x = \sum_t \beta(u,v,w) a_{u,t} = y + y' \quad \text{avec}$$

$$y := \sum_t \gamma(t,u,v,w) a_{u,t} \quad \text{et} \quad y' := \sum_t \gamma'(u,t,v,w) a_{u,t}$$

(on a utilisé 7B6)

Or (cf. 7B4) :

$$y = \sum_t \gamma(u,t,v,w) a_{t,u}$$

et (cf. 7B10 puis 7B9) :

$$y' = \alpha(w) - \beta'(v,u,w)$$

$$y' = \alpha(w) - \sum_t \gamma(v,t,u,w) a_{t,u}$$

On a donc

$$y+y' = \alpha(w) + \sum_t z_t a_{t,u} \quad \text{avec}$$

$$z_t := \gamma(u,t,v,w) - \gamma(v,t,u,w)$$

ce qui implique (cf. 7B7) :

$$z_t = \beta(t,v,w) - \beta(t,u,w)$$

d'où :

$$y+y' = \alpha(w) + \sum_t \beta(t,v,w) a_{t,u} - \sum_t \beta(t,u,w) a_{t,u}$$

c.q.f.d

C. Cas avec second membre

C.1 Théorème fondamental

Soit E un ensemble fini qui comprend au moins trois éléments. Soit a une fonction (réelle ou complexe) définie sur $(E \times E)$ telle que, pour tout élément u de E, $a(u,u) = 0$. Soit x une fonction (réelle ou complexe) définie sur E. On pose :

$$(7C1) \quad x' := \sum_u x(u)$$

$$(7C2) \quad g_{u,w} := \sum_v x_v \beta_{u,v,w}$$

$$(7C3) \quad y_u := \sum_w g_{u,w}$$

$$(7C4) \quad \delta := \sum_u \alpha_u$$

$$(7C5) \quad \eta_u := \delta - \sum_w \lambda_{u,w}$$

On a alors :

$$(7C6) \quad g_{u,w} \sum_t a_{u,t} = \sum_t g_{t,w} a_{t,u} - \alpha_w x_u + (\alpha_w - \lambda_{u,w}) x'$$

et

$$(7C7) \quad y_u \sum_t a_{u,t} = \sum_t y_t a_{t,u} - x_u \delta + x' \eta_u$$

Preuve :

Les éléments u et w étant fixés, on pose :

$$f := g_{u,w} \sum_t a_{u,t} = \sum_{t, v, v \neq u} x_v \beta_{u,v,w} a_{u,t}$$

La relation 7B12 donne :

$$f = \sum_{v, v \neq u} x_v \left\{ \sum_t \beta_{t,v,w} a_{t,u} + \alpha_w - \lambda_{u,w} \right\}$$

$$f = \sum_v x_v \sum_t \beta_{t,v,w} a_{t,u} - x_u \sum_t \beta_{t,u,w} a_{t,u} + (x' - x_u)(\alpha_w - \lambda_{u,w})$$

$$f = \sum_t g_{t,w} a_{t,u} + (x' - x_u) \alpha_w - x' \lambda_{u,w}$$

(par définition de $\lambda_{u,w}$: cf. B.5)

ce qui prouve la relation 7.C.6.

En sommant ces relations sur w , on obtient la relation 7.C.7.

C.2 Cas $x' = 0$: précision

Soit a' une fonction (réelle ou complexe) définie sur $(E \times E)$ (c'est à dire une matrice indexée par $(E \times E)$). Soit a la matrice définie par $a_{u,v} := a'_{u,v}$ si $u \neq v$ et $a_{u,u} := 0$. Soit d la matrice diagonale définie par

$$d_{u,u} := \sum_v a_{u,v} \text{ et } d_{u,v} := 0 \text{ si } u \neq v$$

Soit s la matrice unicolonne (indexée par E) dont tous les termes sont égaux à 1. On suppose que la matrice $a's$ est nulle. Il revient au même de dire que $a' = a - d$. On se propose d'étudier le système suivant

$$(7C8) \quad z a' = x \text{ et } z s = 1$$

où z est la matrice uniligne des inconnues et x une matrice uniligne donnée telle que $x' := x s = 0$.

On considère aussi le système suivant (cf. 2.D. ou 5.A.1) :

$$(7C9) \quad q a' = 0 \text{ et } q s = 1$$

où q est la matrice uniligne des inconnues.

On suppose que le système 7C9 a une seule solution (ce qui est le cas si la matrice a est positive et irréductible : cf. 2.D.3). dans ce cas, soit z et z' deux solutions de 7C8. Si on pose $p := (z - z')$, on a $pa' = 0$ et $ps = 0$ donc $p = 0$. Autrement dit 7C8 n'admet qu'une seule solution.

Soit z la solution de 7C8 et q la solution de 7C9. On suppose $\delta \neq 0$ (ce qui est le cas si a est une matrice positive irréductible). On pose $y := \delta(z - q)$. La matrice uniligne y est telle que $ya' = \delta x$ et $ys = 0$. Ce système ayant une seule solution, le théorème C.1 (relation 7C7) qui précède implique (ici $x' = x s = 0$) que

$$(7C10) \quad y_u = \sum_v \sum_w x_v \beta_{u,v,w}$$

Ceci peut être utilisé sous la forme suivante : q est la solution théorique (par exemple la probabilité stationnaire) que l'on cherche. Soit z une solution approchée (normalisée) obtenue à l'aide d'un algorithme quelconque R .

En général, il est facile de calculer $x = z a'$ (on reporte z dans les équations d'équilibre). La famille de relation 7C10 nous donne l'erreur $z - q = y/\delta$ en fonction de x .

La famille des coefficients $v_{v,u} := \sum_w \beta_{u,v,w} / \delta$ permet donc d'évaluer la précision de l'algorithme R .

On suggère au lecteur d'appliquer ce qui précède pour de petites matrices ($\text{card}(E) = 3$ ou 4) : il pourra constater que le système 7C10 n'est pas évident à obtenir par les méthodes de résolution élémentaires (puisque $xs = 0$, l'écriture de y en fonction de x n'est pas unique).

Evidemment, dans ce qui précède, la solution exacte cherchée q vaut $q = z - y/\delta$: il faut noter cependant que les coefficients β sont plus difficiles à calculer que les coefficients α (qui donnent la solution exacte : cf. A.4). Par contre, on peut espérer obtenir une majoration de β . De plus, il arrive souvent que l'algorithme R (évoqué plus haut) soit tel que x est très petit sauf pour une équation (la "dernière") ; l'essentiel est de majorer les termes associés de la matrice de terme général :

$$\sum_w \beta_{u,v,w}$$

C.3 Lien avec les déterminants

Dans [BKP], le théorème C.1 a été prouvé quand $x' = 0$ par une méthode complètement différente utilisant intensivement les propriétés des déterminants. La preuve initiale du théorème A.4 (cf. [BoM]) reposait, elle aussi, sur les propriétés des déterminants (cf. aussi [Che]).

En un certain sens, les sommes de produits associés aux réseaux fluviaux (coefficients α , β , etc...) introduits aux paragraphes A et B jouent un rôle analogue à celui des déterminants. **Le point essentiel est que, si la matrice a est positive, les quantités α et β ne font intervenir que des sommes de produits de termes positifs : les termes "qui se simplifient" ont été éliminés** (en utilisant le fait que $a's = (a-d)s = 0$) ; attention toutefois, en observant ce qui se passe pour des matrices de petites dimensions, on peut constater que cette élimination est loin d'être triviale.

D. Cas tridiagonal par blocs

D.1 Introduction

On a déjà noté à plusieurs reprises que l'on avait souvent, notamment quand on étudie la probabilité stationnaire associée à un système markovien, une matrice d'évolution a tridiagonale par blocs. Dans ce paragraphe D, on étudie – dans ce cas tridiagonal par blocs – la solution α du système $\alpha(a-d) = 0$ (cas "sans second membre") qui a été mise en évidence dans le théorème A.4.

On suppose donc que E est un ensemble fini, que a est une fonction réelle définie sur $(E \times E)$ avec $a(u,u) := 0$ pour tout élément u de E . La fonction α est définie comme en A.4. On introduit aussi les notations suivantes :

Si p est une fonction définie sur F , partie de E , et à valeurs dans E , on pose :

$$\pi(a, p, F) := \prod_{u \in Z} a(u, p(u))$$

avec $Z := \{u : u \in F \text{ et } u \neq p(u)\}$

Si $F = E$, on a défini en A.3 :

$$\pi(a, r) = \pi(a, r, E)$$

Les hypothèses liées au fait que a est tridiagonale par blocs peuvent être présentées de la façon suivante, qui sera adoptée dans la suite du paragraphe :

Soit n_1 et n_2 deux entiers relatifs avec $n_1 < n_2$. On pose

$K := \{n_1 \leq k \leq n_2\}$. Soit m une application définie sur E et à valeurs dans K . On suppose que $|m(e) - m(e')| > 1$ implique $a(e, e') = 0$. Pour tout élément k de K , on pose $E_k := \{e : m(e) = k\}$. La famille $(E_k)_{k \in K}$ est une partition de E : on a donc, notamment, pour tout réseau fluvial r :

$$\pi(a, r) := \prod_{k \in K} \pi(a, r, E_k)$$

D.2 Notations W et $B(e, w)$

Pour tout élément k de K tel que $n_1 < k < n_2$, on désigne par W'_k (resp. W''_k) l'ensemble des applications définies sur E_k et à valeurs dans E_{k+1} (resp. E_{k-1}). On définit de même W'_k pour $k = n_1$ et W''_k pour $k = n_2$. Par commodité de convention, pour $k = n_1$ (resp. $k = n_2$) on conviendra que W'_{k-1} (resp. W''_{k+1}) n'a qu'un seul élément (qui n'aura donc pas de rôle explicite).

Pour tout élément e de E , on désigne par $W(e)$ l'ensemble des séquences $w := (w_k)_{k \in K'}$ - avec $K' := K \setminus \{m(e)\}$ - telles que, pour $k < m(e)$, w_k appartient à W'_k et, pour $k > m(e)$, w_k appartient à W''_k .

Pour tout élément e de E et pour tout élément w de $W(e)$, on désigne par $G(e, w)$ l'ensemble des réseaux fluviaux r de mer réduite au point e et qui sont "compatibles" avec w au sens suivant :

quels que soient $k < m(e)$ (resp. $k > m(e)$), et u élément de E_k , on a :

$$w_k(u) := r^j(u) \quad \text{avec}$$

$$j := \inf. \{i : r^i(u) \in E_{k+1} \text{ (resp. } E_{k-1})\}$$

quand w parcourt l'ensemble $W(e)$ (e étant fixé), la famille $\{G(e,w)\}_{w \in W(e)}$ constitue une partition de $A(e)$ – qui est l'ensemble des réseaux fluviaux de mer e (cf. A.4). On a donc :

$$\alpha(e) = \sum_{w \in W(e)} \sum_{r \in G(e,w)} \pi(a,r)$$

D.3 Théorème

Pour la commodité du lecteur, rappelons les hypothèses : **E est un ensemble fini, n_1 et n_2 sont deux entiers relatifs avec $n_1 < n_2$, m est une application définie sur E et à valeurs dans l'ensemble K des entiers relatifs k tels que $n_1 \leq k \leq n_2$, et a est une application définie sur $(E \times E)$ telle que, quel que soit u élément de E , $a(u,u) = 0$.**

Soit d la matrice diagonale définie par

$$d(u,v) := \delta_{u,v} \sum_{w \in E} a(u,w)$$

et α la solution du système $\alpha(a-d) = 0$ construite en A.4. On suppose que $|m(u)-m(v)| > 1$ implique $a(u,v) = 0$. On pose $E_k := \{u : m(u) = k\}$.

Alors il existe des matrices $(S_k)_{k \in K}$ et $(T_k)_{k \in K}$ et, pour tout élément e de E , $(R_{k,e})_{k \in K}$ dont les dimensions et les valeurs seront explicitées au cours de la démonstration et telles que les propriétés suivantes soient satisfaites :

(i) **S_k , T_k et $R_{k,e}$ ne dépendent que de la restriction de a à l'ensemble $(E_k \times E)$.**

(ii) **pour $n_1 < k < n_2$ et e élément de E_k , on a :**

$$(7D1) \quad \alpha(e) = S_{n(1)} \dots S_j S_{j+1} \dots S_{k-1} R_{k,e} T_{k+1} \dots T_h T_{h+1} \dots T_{n(2)}$$

De plus, si la fonction a est positive, les matrices S_k , T_k et $R_{k,e}$ sont positives.

Preuve :

1°) On utilise les notations introduites en D.1 et D.2. Pour tout élément e de E et tout élément w de $W(e)$, on pose :

$$(7D2) \quad \theta(e,w) := \sum_{r \in G(E,w)} \prod_{k \in K} \pi(a,r,E_k)$$

Les dernières lignes de D.1 et D.2 respectivement donnent :

$$(7D3) \quad \alpha(e) = \sum_{w \in W(c)} \theta(e,w)$$

On va maintenant transformer le second membre de 7.D.2.

2°) Pour $n_1 < k < n_2$ on pose $V_k := E_{k-1} \cup E_k \cup E_{k+1}$. Pour $k = n_1$ on pose $V_k := E_k \cup E_{k+1}$ et, pour $k = n_2$, on pose $V_k := E_{k-1} \cup E_k$.

On désigne par P_k (resp. Q_k) l'ensemble des fonctions définies sur E_k (resp. V_k) et à valeurs dans V_k .

3°) Quels que soient k élément de K et x élément de W'_{k-1} , $\psi'(k,x)$ est une fonction définie comme suit : $\psi'(k,x)$ est une application définie sur P_k et à valeurs dans Q_k ; plus précisément, si p est un élément de P_k ,

$q := \psi'(k,x)(p)$ est un élément de Q_k défini comme suit :

- si u appartient à E_k , $q(u) := p(u)$
- si u appartient à E_{k+1} , $q(u) := u$
- si $k > n_1$ et u appartient à E_{k-1} , $q(u) := x(u)$

(si $k = n_1$, W'_{k-1} n'a qu'un seul élément x et $\psi'(k,x)$ ne dépend pas de x).

On définit $\psi''(k,z)$ pour z élément de W''_{k+1} de façon analogue en intervertissant les rôles de $(k-1)$ et $(k+1)$.

4°) Quels que soient k élément de K et (x,y) élément de $(W'_{k-1} \times W'_k)$, on appelle $F'(k,x,y)$ l'ensemble des éléments p de P_k tels que, si on pose $q := \psi'(k,x)(p)$, q est un réseau fluvial (propre) (sur V_k) et pour tout élément u de E_k , $q^*(u) = y(u)$ (cf. A.3 pour la définition de q^*).

On définit de même $F''(k,x,y)$ pour (x,y) élément de $(W''_{k+1} \times W''_k)$ en intervertissant les rôles de $(k-1)$ et $(k+1)$.

5°) On pose :

$$S_k(x,y) := \sum_{p \in F'(k,x,y)} \pi(a,p,E_k)$$

$$T_k(x,y) := \sum_{p \in F'''(k,x,y)} \pi(a,p,E_k)$$

6°) Quels que soient k élément de K et (x,z) élément de $(W'_{k-1} \times W''_{k+1})$, $\rho(k,x,z)$ est une fonction définie sur P_k et à valeurs dans Q_k ; plus précisément, si p est un élément de P_k , $q := \rho(k,x,z)(p)$ est l'élément de Q_k défini comme suit :

- si u appartient à E_k , $q(u) := p(u)$
- si u appartient à E_{k-1} , $q(u) := x(u)$
- si u appartient à E_{k+1} , $q(u) := z(u)$

De plus, si e est un élément de E tel que $m(e) = k$, on désigne par $H(e,x,z)$ l'ensemble des éléments p de P_k tels que, si on pose $q := \rho(m(e),x,z)$, q est un réseau fluvial (propre) (sur V_k) de mer e .

Enfin, on pose :

$$R_{k,e}(x,z) := \sum_{p \in H(e,x,z)} \pi(a,p,E_k)$$

7°) Quels que soient e élément de E , w élément de $W(e)$ et k élément de K , on définit $F(k,e,w)$ de la façon suivante :

- pour $k < m(e)$, $F(k,e,w) := F'(k, w_{k-1}, w_k)$
- pour $k = m(e)$, $F(k,e,w) := H(e, w_{k-1}, w_{k+1})$
- pour $k > m(e)$, $F(k,e,w) := F'''(k, w_{k+1}, w_k)$

8°) Pour tout élément r de $G(e,w)$ et tout élément k de K , soit r_k la restriction de r à E_k . On peut donc écrire $r := (r_{n(1)}, \dots, r_k, \dots, r_{n(2)})$.

L'étape essentielle de la démonstration est de noter que r appartient à $G(e,w)$ si et seulement si, quel que soit k élément de E_k , r_k appartient à $F(k,e,w)$. La relation 7D2 implique donc

$$(7D4) \quad \theta(e,w) = \prod_{k \in K} \sum_{p \in F(k,e,w)} \pi(a,p,E_k)$$

9°) Compte tenu de 7D3 et 7D4, la relation 7D1 se vérifie alors simplement en développant les produits matriciels.

D.4 Commentaires heuristiques

1°) La preuve proposée précédemment est complète du point de vue formel théorique. Par contre elle est assez difficile à assimiler. Disons d'abord qu'une étape essentielle pour comprendre cette preuve est d'observer, à l'aide de dessins, les divers cas possibles, en commençant par le cas où E_k a deux ou trois éléments.

2°) Pour faciliter ces commentaires "heuristiques", on suppose que les ensembles E_k sont dessinés sur un plan vertical (un "tableau noir"), à chaque ensemble E_k correspond une ligne verticale, les lignes les plus à gauche correspondant aux plus petites valeurs de k .

3°) On fixe k . Les éléments p de $F'(k,x,y)$ sont, en simplifiant, les restrictions à E_k des réseaux fluviaux "compatibles" avec x et y ; ils n'interviennent que si e est à droite de k .

L'élément x de W_{k-1}' correspond à l'information minimale qu'il faut transmettre de $(k-1)$ à k pour décider si p est associé ou non à un réseau fluvial. De même, l'élément y de W_k' correspond à l'information qu'il faut transmettre pour l'étape suivante.

4°) On a donc une situation de type "markovien" en espace mais dans un cadre tout à fait nouveau. Quand on va de la gauche vers la droite, ce sont les éléments de W_k' qui donnent l'information minimale pour avoir une évolution markovienne quand k augmente. Au contraire, quand on va de la droite vers la gauche (k diminue) il faut considérer les éléments de W_k'' .

D.5 Remarques "techniques"

La présentation proposée en D.3 est la plus simple du point de vue théorique ; par contre, on peut lui apporter beaucoup d'améliorations "techniques" (cf. [Pel.7]).

1°) Les seuls réseaux fluviaux r qui interviennent effectivement dans le calcul de α sont ceux pour lesquels u diffère de $r(u)$ implique $a(u,r(u))$ différent de zéro : ceci est évidemment **fondamental** du point de vue technique.

2°) Notamment, pour tout élément k de K , soit E'_k (resp. E''_k) l'ensemble des éléments u de E_k pour lesquels il existe v élément de E_{k+1} (resp. E_{k-1}) avec $a(v,u)$ différent de zéro. Dans tout ce qui précède, au lieu de définir W'_k comme cela a été fait en D.2, **on peut prendre pour W'_k l'ensemble des applications définies sur E'_k (et non E_k) et à valeurs dans E''_{k+1}** . Par la suite, les autres notations doivent être modifiées en conséquence. Evidemment on peut, de même, définir W''_k comme étant l'ensemble des applications définies sur E''_k et à valeurs dans E'_{k-1} .

3°) Dans le cas particulier "avec pivot", quel que soit k , E'_k (ou E''_k suivant les conventions utilisées) **n'a qu'un seul élément** : W''_k n'a donc qu'un seul élément et les matrices T_k sont de dimensions (1×1) (cf. [BeP]).

4°) Pour alléger les notations, on a utilisé le même symbole S quand $k > n_1$ et quand $k = n_1$ (matrice "terminale") ; en pratique n_1 et/ou n_2 **varient** en fonction de la précision et des valeurs des paramètres : il y a donc intérêt à choisir un symbole spécifique (par exemple $S'_{n(1)}$) au bord pour que les matrices S_k ne dépendent pas de n_1 et n_2 .

5°) **Les dimensions des matrices S_k , etc... peuvent varier avec k** ; ce peut être le cas même si k est le nombre de clients dans une station. Il ne faut pas oublier qu'il y a un très grand nombre de façons d'obtenir une situation tridiagonale par blocs.

6°) Dans la relation 7D1 du théorème D.3 les matrices S_j et T_h sont, en général, **bien mélangeantes** en ce sens que leur coefficient de contraction (cf. 2.A.3) est nettement inférieur à 1. La relation 7D1 montre donc que, si $n_2 - n_1$ est assez grand, les valeurs de $\alpha(e)$ pour e élément de $E_{n(1)}$ dépendent relativement peu des "contraintes" imposées en n_2 (matrice $T_{n(2)}$) ; **la relation 7D1 est donc la réponse "théorique" aux questions posées en 5.D.4** ; évidemment, on peut utiliser l'algorithme proposé en 5.D même si on n'a pas compris le théorème 7.D.3 !

7°) Il ne faut pas confondre la méthode ici proposée avec l'utilisation des "probabilités taboues" (cf., par exemple, [Chu]) qui est reprise dans [Neu] (cf. aussi [Wal]) dans le cas particulier des matrices fixes.

D.6 Cas E infini

En général, dans ce cours, pour faciliter l'étude théorique, on s'est restreint au cas où E est un ensemble fini. Or un très grand nombre d'études théoriques portent sur le cas où E est infini : étude des états récurrents, transients, etc... (cf. par exemple, [Rev]). En fait, autant que l'auteur ait pu en juger, ces travaux théoriques prestigieux **ne permettent pas de répondre aux questions posées en 5.D.4.**

Inversement, le théorème D.3 donne des conditions suffisantes pour que les probabilités d'états en régime stationnaire à "un bord" - qui peut être le "centre" - dépendent relativement peu de "ce qui se passe à l'autre bord" - qui peut être le "voisinage de l'infini" : ceci donne donc une technique totalement nouvelle pour aborder le cas où E est infini.

E. Liens entre homogène et non homogène

E.1 Introduction

Le but de ce paragraphe est de préciser certaines relations entre la résolution des systèmes homogènes et celle des systèmes non homogènes. On notera aussi les éclairages complémentaires donnés en 6.A.4, 6.C.2 et, surtout, au paragraphe C de ce chapitre. Enfin, il y a évidemment une très grande analogie entre ce qui suit et l'étude des systèmes différentiels linéaires, homogènes et non homogènes. Notamment, le théorème E.3 qui suit montre comment on peut ramener partiellement la résolution d'un système non homogène à celle d'un système homogène (cf. E.4).

E.2 Drainage

Définition : Etant donné un ensemble fini E, une matrice positive a indexée par $(E \times E)$ et F une partie de E, on dira que E est a-drainé par F si, pour tout élément u de E il existe un élément v de F tel que v est a-irrigable par u.

Rappelons que la définition de l'irrigabilité a été donnée en A.5. Notons aussi que, dans la proposition A.6, l'hypothèse $G=E$ équivaut à dire que E est a-drainé par F. Le théorème qui suit est, en un certain sens, une généralisation de cette proposition A.6.

E.3 Théorème de base

Soit E un ensemble fini et (H, K) une partition de E , H et K étant non vides. Soit d et x deux fonctions positives définies sur E et b une fonction positive définie sur $(E \times E)$. On suppose que, pour tout élément e de E , on a $b(e, e) = 0$ et

$$(7E1) \quad d(e) \geq \sum_{e' \in E} b(e, e')$$

et que pour tout élément i de H on a :

$$(7E2) \quad \sum_{e \in K} b(i, e) > 0$$

Soit F l'ensemble des éléments e de K tels que l'inégalité 7E1 est une inégalité stricte, c'est à dire que

$$(7E3) \quad d(e) > \sum_{e' \in E} b(e, e')$$

On suppose que E est b -drainé par F . On a alors :

1°) Pour tout élément i de H , il existe une fonction positive q_i définie sur K telle que, pour tout élément e de K , on a :

$$(7E4) \quad q_i(e) d(e) = \sum_{e' \in K} q_i(e') b(e', e) + b(i, e)$$

2°) Il existe une fonction positive q' définie sur K telle que, pour tout élément e de K , on a :

$$(7E5) \quad q'(e) d(e) = \sum_{e' \in K} q'(e') b(e', e) + x(e)$$

3°) Il existe une fonction u définie sur H telle que, si on pose :

$$\text{pour } e \text{ élément de } K, q(e) := q'(e) + \sum_{i \in H} u_i q_i(e)$$

$$\text{pour } i \text{ élément de } H, q(i) := u_i$$

alors, pour tout élément e de E , on a :

$$(7E6) \quad q(e) d(e) = \sum_{e' \in E} q(e') b(e', e) + x(e)$$

Preuve :

1°) On considère donc un élément i fixé de H . Soit K' l'ensemble K auquel on a adjoint un point extérieur que l'on note 0 (donc 0 n'appartient pas à K). Soit d' la fonction positive définie sur K' par $d'(e) := d(e)$ si e est différent de 0 et

$$d'(0) := \sum_{e \in K} b(i, e)$$

De même, soit b' la fonction positive définie sur $(K' \times K')$ par :

$$b'(e, e') := b(e, e') \text{ si } e \text{ et } e' \text{ appartiennent à } K$$

$$b'(e, 0) := 0 \text{ pour tout élément } e \text{ de } K'$$

$$b'(0, e) := b(i, e) \text{ si } e \text{ appartient à } K$$

$$b'(e, 0) := d(e) - \sum_{e' \in K} b(e, e') \text{ si } e \text{ appartient à } K$$

D'une part, pour tout élément e de K' , on a

$$d'(e) = \sum_{e' \in K'} b'(e, e')$$

D'autre part, 0 est b' -irrigable par n'importe quel élément e de K (compte tenu de la définition de $b'(e, 0)$ et puisque E est b -drainé par F). Il existe donc (cf. A.4) une fonction positive q_i définie sur K' telle que $q_i(0) = 1$ et telle que, pour tout élément e de K' on ait :

$$(7E7) \quad q_i(e) d'(e) = \sum_{e' \in K'} q_i(e') b'(e', e)$$

Ceci implique la relation 7E4.

2°) La preuve de ce 2°) est tout à fait analogue à celle du 1°). On définit K' comme au 1°). Soit d' la fonction positive définie sur K' par $d'(e) := d(e)$ si e appartient à K et

$$d'(0) := \sum_{e \in K} x(e)$$

Soit b' la fonction positive définie sur $(K' \times K')$ par :

$$\text{pour tout élément } e \text{ de } K', b'(e, e) := 0$$

$$\text{pour tout couple } (e, e') \text{ d'éléments de } K, b'(e, e') := b(e, e')$$

pour tout élément e de K , $b'(0,e) := x(e)$ et

$$b'(e,0) := d(e) - \sum_{e' \in K} b(e,e')$$

D'une part, pour tout élément e de K' , on a

$$d'(e) = \sum_{e' \in K'} b'(e,e')$$

D'autre part, 0 est b' -irrigable par n'importe quel élément e de K (pour les mêmes motifs qu'au 1°). Il existe donc une fonction positive q' définie sur K' telle que $q'(0) = 1$ et telle que, pour tout élément e de K' , on a :

$$q'(e) d'(e) = \sum_{e' \in K'} q'(e') b'(e',e)$$

ce qui implique la relation 7E5.

3°) Compte tenu de la définition de q et des relations 7E4 et 7E5, on vérifie alors immédiatement que, quelle que soit la fonction u , la relation 7E6 est satisfaite pour tous les éléments e de K .

Considérons ensuite un élément i de H . La relation 7E6 (qu'il faut démontrer) s'écrit :

$$\begin{aligned} (7E8) \quad u_i d(i) &= \sum_{j \in H} u_j b(j,i) + x(i) + \sum_{e \in K} q(e) b(e,i) \\ &= y(i) + \sum_{j \in H} u_j c(j,i) \quad \text{en ayant posé :} \\ y(i) &:= x(i) + \sum_{e \in K} q(e) b(e,i) \quad \text{et} \\ c(j,i) &:= b(j,i) + \sum_{e \in K} q_j(e) b(e,i) \end{aligned}$$

Soit d'' la "matrice" indexée par $(H \times H)$ et définie par $d''(i,i) := d(i)$ et $d''(i,j) := 0$ si $i \neq j$. La relation 7E8 peut s'écrire matriciellement $u d'' = y + u c$ soit $u(d'' - c) = y$. Nous allons maintenant prouver que $(d'' - c)$ est une M -matrice inversible, ce qui montrera l'existence de la fonction positive u et achèvera la preuve du théorème.

4°) Par construction, $(d'' - c)$ est une L -matrice (cf. 5.A.1.-4°). Etudions la somme des colonnes de c . On a :

$$(7E9) \quad \sum_{j \in H} c(i,j) = \sum_{j \in H} b(i,j) + \sum_{j \in H} \sum_{e \in K} q_i(e) b(e,j)$$

La relation 7E1 peut s'écrire

$$\sum_{j \in H} b(e,j) \leq d(e) - \sum_{e' \in K} b(e,e')$$

ce qui implique

$$(7E10) \quad \sum_{e \in K} \sum_{j \in H} q_i(e) b(e,j) \leq \sum_{e \in K} q_i(e) d(e) - \sum_{e \in K} \sum_{e' \in K} q_i(e) b(e,e')$$

soit, compte tenu de 7E4 :

$$\sum_{e \in K} \sum_{j \in H} q_i(e) b(e,j) \leq \sum_{e \in K} b(i,e)$$

ce qui reporté dans 7E9 donne :

$$\begin{aligned} \sum_{j \in H} c(i,j) &\leq \sum_{j \in H} b(i,j) + \sum_{e \in K} b(i,e) \\ &\leq \sum_{e \in E} b(i,e) \end{aligned}$$

La relation 7E1 implique alors

$$(7E11) \quad \sum_{j \in H} c(i,j) \leq d(i)$$

ce qui montre que $(d''-c)$ est une L-matrice à diagonale dominante au sens large.

5°) Soit i un élément fixé de H . Soit $G(i)$ l'ensemble des éléments e de K tels que $q_i(e) > 0$; cet ensemble n'est pas vide (cf. 7E2 et 7E4). Puisque K est b -drainé (et donc b' -drainé) par F , il existe un élément e qui appartient à la fois à F et à $G(i)$ (cf. le 1°) de la preuve). On a donc à la fois $q_i(e) > 0$ et 7E1 est une inégalité stricte ; 7E10 est donc aussi une inégalité stricte ainsi que 7E11. La matrice $(d''-c)$ est donc une L-matrice à diagonale strictement dominante. C'est donc une M-matrice inversible (résultat classique qui est un cas particulier de 7.A.6).

E.4 Autre présentation

On garde les notations données dans le théorème E.3 qui précède. Considérons le système 7E6, soit

$$(7E6) \quad q(e) d(e) = \sum_{e' \in E} q(e') b(e',e) + x(e)$$

la famille des "inconnues" étant la famille q (indexée par E).

La solution q' de 7E5 correspond alors à la solution particulière de 7E6 dans laquelle on impose $q(e) = 0$ pour e élément de H . De même, la solution q_i de 7E4 correspond, à une constante multiplicative près, à la solution du système homogène associé à 7E6 (c'est à dire que $x = 0$) et où on a imposé $q_i(i) = 1$ et $q_i(e) = 0$ si e appartient à H , $e \neq i$. Il peut d'ailleurs être opportun de choisir d'autres contraintes que celles-là.

En langage imagé, on peut dire que **la solution générale q "avec second membre" est la somme d'une solution particulière q' avec second membre et de la "solution générale" $\sum u_i q_i$ du système sans second membre.**

Par ailleurs, cette présentation montre que pour résoudre les systèmes 7E4 et 7E5, on peut utiliser **un seul et même algorithme**, seules les "initialisations" de $q(i)$ pour i élément de H et de x changent : cette technique a été abondamment utilisée par l'auteur et ses collaborateurs (cf. [Oum], [Khe], [Sag], etc...) ; la même remarque avait été effectuée en 5.B.5. De plus, comme cela a été souvent noté dans ce cours, si on utilise cette technique les divers calculs peuvent être menés en parallèle pour les diverses valeurs de i .

E.5 Remarques

1°) Le théorème 7E3 donne aussi une méthode très puissante de **décomposition** de certains systèmes linéaires ; en effet, puisque K est une partie de E , la restriction de b à K peut éventuellement être "décomposable" ou "presque décomposable" : il suffit que les "liaisons" entre divers sous-ensembles de K s'effectuent uniquement par les éléments de H .

2°) Rappelons qu'il est parfois intéressant de ramener la résolution d'un système non homogène à n inconnues à celle d'un système homogène à $(n+1)$ inconnues : on procède comme dans le théorème A.6.

3°) Il y a beaucoup d'autres présentations possibles du théorème E.3 : notamment on peut le prouver à partir de "contraintes artificielles" comme en 5.E.3.6°).

4°) Evidemment l'énoncé et la preuve du théorème E.3 ont une (ou plusieurs) interprétation "probabiliste" de même que 6.C.2 était une interprétation probabiliste de 6.A.4.

5°) Le théorème E.3 peut être intéressant même si H n'a qu'un seul élément (cf. 5.B.4).

Chapitre 8

Régime transitoire**A Exponentielle de matrice****A.1 Introduction**

On a déjà amorcé, au chapitre 1, l'étude du régime transitoire, c'est à dire l'étude de la "loi du processus" (cf. 1.B. notamment). En fait, cette loi ne peut être effectivement explicitement calculée que pour des processus possédant des propriétés extrêmement restrictives ou ne comportant que peu d'états - relativement aux moyens de calculs utilisés.

Par contre, comme cela a été noté en 1.B.4, si le processus considéré est markovien homogène et si l'ensemble de ses états est fini, la loi de ce processus est "**théoriquement**" complètement déterminée par la **loi initiale** et le **semi-groupe de transition** ; de plus, la forme théorique de ce semi-groupe est connue (cf. 1.D.5) : **c'est une exponentielle de matrice**, solution du système différentiel linéaire des équations de Kolmogorov.

Plus généralement, il y a une relation très étroite entre la notion d'exponentielle de matrices et les solutions d'un système différentiel linéaire à coefficients constants : cette relation est connue depuis très longtemps mais son intérêt est relancé par l'utilisation des moyens informatiques. En effet, autant que l'auteur ait pu en juger, **les résultats les plus exploitables numériquement - qui vont être explicités dans ce chapitre 8 - sont les résultats les plus simples reposant sur les propriétés de base des exponentielles de matrice.**

Pour la commodité du lecteur, nous allons reprendre complètement cette étude en résumant toutefois les points que l'on peut trouver dans n'importe quel cours d'analyse élémentaire. Nous allons nous limiter au cas où l'ensemble E des états est fini : le cas où E est infini dénombrable peut être abordé comme cela a été indiqué en 1.D.5.

A.2 Propriétés "élémentaires"

On se donne un ensemble fini E et une fonction m indépendante de t définie sur $(E \times E)$; cette fonction m peut être considérée comme une matrice m (en un sens "généralisé" : cf. 5.A.1) indexé par $(E \times E)$.

Théorème : (cf. 1.D.4)

Il existe une et une seule fonction définie sur l'axe réel et à valeurs dans l'ensemble des matrices indexées par $(E \times E)$ - dont la valeur en t sera

notée e^{mt} - telle que les propriétés suivantes soient satisfaites :

$$(8A1) \quad \frac{d}{dt}(e^{mt}) = m e^{mt} = e^{mt} m$$

$$(8A2) \quad e^{m0} \text{ est la matrice unité}$$

$$(8A3) \quad \text{pour tout couple (s,t) de réels,}$$

$$e^{m(s+t)} = e^{mt} e^{ms}$$

De plus, on a :

$$(8A4) \quad (e^{mt})_{u,v} = \sum_{k=0}^{\infty} (m^k)_{u,v} t^k / k !$$

cette série entière ayant un rayon de convergence infini (avec la convention : m^0 est la matrice unité).

Preuve :

Soit i la matrice unité (indexée par $(E \times E)$).

$$1^\circ) \text{ On pose } d := \sup_{u \in E} \sum_{v \in E} |(m)_{u,v}|$$

On vérifie, par récurrence croissante sur k , que $\sum_{v \in E} |(m^k)_{u,v}| \leq d^k$

et la série qui apparaît dans le deuxième membre de 8A4 est majorée terme à terme par $(d|t|)^k / k !$ ce qui est le terme général d'une série numérique convergente (la série exponentielle classique) ; il y a donc convergence normale et le rayon de convergence est infini. On peut donc définir $(e^{mt})_{u,v}$ à partir de 8A4 ; on peut aussi dériver terme à terme d'où on déduit facilement 8A1. La relation 8A2 découle de 8.A.4.

2°) Soit $(h(t))_{t \in \mathbb{R}}$ une famille de matrices telle que l'on ait à la fois :

$$(i) \quad \frac{d}{dt}(h(t)) = m h(t)$$

et

$$(ii) \quad h(0) \text{ est la matrice unité.}$$

On pose $g(t) = e^{-mt} h(t)$

$$\text{On a } \frac{d}{dt}(g(t)) = \frac{d}{dt}(e^{-mt}) \cdot h(t) + e^{-mt} \cdot \frac{d}{dt}(h(t))$$

(propriété classique et facile à vérifier de la dérivation d'un produit matriciel)

$$\frac{d}{dt}(g(t)) = -e^{-mt} m h(t) + e^{-mt} m h(t) = 0$$

La matrice $g(t)$ est donc indépendante de t et $g(t) = g(0) = i$ est la matrice unité.

3°) Si on prend $h(t) = e^{mt}$, on déduit du 2°) ci-dessus que e^{-mt} est la matrice inverse de la matrice e^{mt} .

4°) Si $h(t)$ est une matrice "quelconque" satisfaisant aux conditions 2°) (i) et (ii), le 2°) implique que $e^{-mt} h(t) = i$ soit, en utilisant le 3°), $h(t) = e^{mt}$: autrement dit les conditions (i) et (ii) caractérisent la matrice e^{mt} .

5°) Si on prend $h(t) = e^{m(s+t)} e^{-ms}$ on a $\frac{d}{dt}(h(t)) = m h(t)$ et $h(0) = e^{ms} e^{-ms} = i$ donc $h(t) = e^{mt}$. On en déduit 8A3.

A.3 Cas commutatif

Soit p et q deux matrices indexées par $(E \times E)$ telles que $p q = q p$. On a alors $e^{(p+q)t} = e^{qt} e^{pt}$.

Preuve :

Soit (s, t) un couple de réels. On raisonne comme au 5°) de A.2 ; soit $h(t) = e^{(ps+qt)} e^{-ps}$. On a $\frac{d}{dt}(h(t)) = q h(t)$ et $h(0) = i$ donc (cf. A.2) $h(t) = e^{qt} = e^{(ps+qt)} e^{-ps}$ d'où $e^{ps+qt} = e^{qt} e^{ps}$, c.q.f.d. (avec $s = t$).

A.4 Cas "positif"

Soit m une matrice indexée par $(E \times E)$ telle que $u \neq v$ implique $m_{u,v} \geq 0$. On a alors, quel que soit $t \geq 0$, $(e^{mt})_{u,v} \geq 0$.

Preuve :

On pose $c := \sup_{u \in E} |(m)_{u,u}|$

D'une part $(m+c i)$ est une matrice positive donc, pour $t \geq 0$, $e^{(m+c i)t}$ est une matrice positive.

D'autre part, les matrices m et c commutent, donc $e^{mt} = e^{-cit} e^{(m+c)i}t$ (cf. A.3). Enfin, on déduit immédiatement de 8A4 que $e^{-cit} = e^{-c} i$. On a donc finalement :

$$(8A5) \quad e^{mt} = e^{-c} t e^{(m+c)i}t$$

Ceci implique notamment que e^{mt} est une matrice positive (pour $t \geq 0$).

La technique consistant à utiliser systématiquement la relation 8A5, et dont l'intérêt a été "redécouvert" récemment, est parfois appelée **uniformisation**.

A.5 Calcul de e^{mt}

1°) D'innombrables techniques ont été proposées pour calculer e^{mt} dans divers cas particuliers. En fait, les méthodes les plus efficaces sont souvent celles associées aux propriétés élémentaires rappelées dans ce paragraphe A. Plus précisément, sauf dans quelques cas particuliers, le calcul général le plus robuste de e^{mt} repose sur les relations 8A3, 8A4 et 8A5 avec des variantes qui dépendent du contexte (cf. [GoV] et [Abd]).

2°) Par exemple, on peut d'abord utiliser la relation 8A5 si la matrice $(m+ci)$ est positive.

3°) Par ailleurs, soit m' une matrice (positive ou non) et soit :

$$d' := \sup_{u \in E} \sum_{v \in E} |(m')_{u,v}|$$

On peut calculer $e^{m't}$ à partir de la relation 8A4 pour t "assez petit", par exemple pour $d't < 1$ (pour avoir une convergence assez rapide de la série exponentielle).

4°) Ensuite, $e^{m't}$ étant calculé, on utilise (cf. 8A3) l'égalité $e^{m's} = (e^{m't})^{f(k)}$, $f(k) := 2^k$, pour $s = t.2^k$, ce qui revient à élever $(e^{m't})$ k fois au carré pour calculer $e^{m's}$.

A.6 Condition initiale

Notons que, au 4°) du paragraphe A.5 qui précède, le calcul du carré de $e^{m't}$ suppose de connaître (c'est à dire d'avoir "en mémoire") toute la matrice $e^{m't}$.

Par contre, au 3°) de ce même paragraphe A.5, si la matrice m' est mise "en mémoire", on peut calculer $e^{m't}$ **ligne par ligne ou colonne par colonne** à partir de la relation 8A4 : en effet, par exemple, pour calculer la

j-ième ligne de $(m')^{k+1}$, il suffit de disposer de la matrice m' et de la j-ième ligne de la matrice $(m')^k$ en utilisant la relation

$$(m')^{k+1} = (m')^k m'$$

D'une part, ceci permet, par exemple, de faire calculer les diverses lignes **par divers processeurs opérant en parallèle** (calcul "vectoriel").

D'autre part, cette remarque est techniquement importante quand on a seulement à calculer $z e^{m't}$ où **z est une matrice uniligne donnée** (ce qui est évidemment beaucoup moins contraignant que de calculer toute la matrice $e^{m't}$) : plus précisément, on utilise les deux relations suivantes :

$$z(m')^{k+1} = [z(m')^k] m'$$

$$\text{et } z e^{m't} = \sum_{k=0}^{\infty} [z(m')^k] t^k / k !$$

Notamment, rappelons (cf. 1.B.4 et 1.D.5) que si X est un processus markovien homogène qui admet E comme ensemble d'états, a comme "matrice d'évolution" et z comme "loi initiale", on a :

$$\text{Proba } [X_t = u] = (z e^{(a-d)t})_u$$

A.7 Remarques

1°) Il y a un très grand nombre de méthodes classiques pour calculer une exponentielle de matrice : notons toutefois que, désormais, on s'intéresse surtout aux méthodes que l'on peut implémenter sur ordinateur.

Par exemple, si la matrice m est triangulaire, le calcul de e^{mt} se ramène à la résolution d'une séquence (indexée par E) d'équations différentielles linéaires avec "second membre", chacune de ces équations comportant une seule inconnue et le second membre étant une combinaison linéaire de termes de la forme $t^k e^{-ut}$: la résolution de ces équations peut alors être effectuée de façon "formelle" (cf. [MRT] et 5.A.1, 5°).

2°) Le lecteur est probablement un peu surpris de voir avec quelle insistance les propriétés élémentaires de l'exponentielle de matrice sont rappelées et détaillées dans ce cours : pourtant, la lecture de diverses communications, mêmes récentes, devraient le convaincre que ce rappel n'est peut-être pas inutile. De plus, un calcul sophistiqué de e^{mt} n'est valable que s'il est plus efficace – ou, du moins, peut devenir plus efficace – que les calculs reposant sur les méthodes élémentaires évoquées ci-dessus.

B Splitting et chaîne incluse

B.1 Exponentielle de $(m+n)$

Théorème : Soit m et n deux matrices (carrées) indexées par $(E \times E)$. On suppose que m et n ne commutent pas ($m n \neq n m$) sinon ce qui suit n'a pratiquement pas d'intérêt (cf. 8A3).

Pour tout couple (k, t) avec k entier et t réel (positif), soit $q_k(t)$ une matrice indexée par $(E \times E)$. On suppose que la série de fonctions matricielles $(q_k(t))_{k \geq 0}$ satisfait aux propriétés suivantes :

- (i) $q_0(o) = i$ et, pour $k > 0$, $q_k(o) = 0$
où i est la matrice unité
- (ii) $\frac{d}{dt} q_0(t) = n q_0(t)$
- (iii) pour $k \geq 0$, $\frac{d}{dt} q_{k+1}(t) = m q_k(t) + n q_{k+1}(t)$

On suppose aussi que les séries matricielles $(q_k(t))_{k \geq 0}$ et $(\frac{d}{dt} q_k(t))_{k \geq 0}$ convergent uniformément (ce qui autorise à dériver terme à terme) pour $t \geq 0$. On a alors :

$$e^{(m+n)t} = \sum_{k=0}^{\infty} q_k(t)$$

Preuve :

On pose $r(t) := \sum_{k=0}^{\infty} q_k(t)$. On vérifie immédiatement que l'on a $r(o) = i$ et $r'(t) = (m+n) r(t)$ d'où le résultat (cf. A.2).

B.2 Remarques

1°) Ce "théorème" est l'analogue, dans le cas transitoire, du "théorème" 6.A.2 associé au cas stationnaire. Là encore, en un certain sens, l'intérêt de ce "théorème" est de "séparer" l'étude associée à m de celle associée à n . Certes, l'essentiel reste à faire au niveau des exemples, c'est à dire trouver une bonne décomposition (un bon splitting) (m, n) de $(m+n)$.

2°) On vérifie immédiatement que l'on a :

$$(8B1) \quad q_0(t) = e^{nt}$$

et, pour $k \geq 0$,

$$(8B2) \quad q_{k+1}(t) = \int_0^t q_0(s) m q_k(t-s) ds = \int_0^t q_0(t-s) m q_k(s) ds$$

B.3 Cadre probabiliste

Reprenons les hypothèses les plus souvent introduites dans ce cours (cf., notamment, 1.D, 2.D et 6.A) : a est une matrice positive indexée par $(E \times E)$ dont la diagonale est nulle (i.e. $a_{u,u} = 0$ pour tout élément u de E) ; d est la matrice diagonale associée à a et définie par

$$d_{u,v} := \delta_{u,v} \sum_{w \in E} a_{u,w}$$

Soit b et c deux matrices positives indexées par $(E \times E)$ telles que $a = b + c$. On pose :

$$c' := \sup_u \sum_v c_{u,v}$$

On pose $n := b - d$ et $m := c$.

On définit la famille $(q_k(t))_{k \geq 0}$ par récurrence croissante sur k à partir des relations 8B1 et 8B2. Rappelons que la matrice $q_0(t)$ est positive (cf. A.4).

Quels que soient u élément de E et $t \geq 0$, on a :

$$\sum_{v \in E} \frac{d}{dt} [q_0(t)]_{u,v} = \sum_{v \in E} [q_0(t)(b-d)]_{u,v} \leq 0 \quad \text{et} \quad \sum_{v \in E} [q_0(0)]_{u,v} = 1$$

ce qui implique $\sum_{v \in E} [q_0(t)]_{u,v} \leq 1$

ou encore :

$$\sup_u \sum_v [q_0(t)]_{u,v} \leq 1$$

On déduit alors de 8B2, (avec $m = c$), en raisonnant par récurrence croissante sur k , que les matrices $q_k(t)$ sont positives et que l'on a :

$$\sup_{u \in E} \sum_{v \in E} [q_{k+1}(t)]_{u,v} \leq (c' t)^k / k !$$

On en déduit que les séries matricielles de terme général $q_k(t)$ et $\frac{d}{dt} q_k(t)$ respectivement sont uniformément convergentes pour $t \geq 0$. On peut donc appliquer le "théorème" B.1.

B.4 Interprétation heuristique

Nous allons maintenant donner l'interprétation "probabiliste" du théorème B.1 sous les hypothèses proposées au paragraphe B.3 ci-dessus. Une justification complète et rigoureuse du théorème B.1 à partir de l'étude des processus dépasserait le cadre de ce cours : comme en 6.C est-il d'ailleurs nécessaire de donner une telle preuve alors que le théorème à prouver est quasi-trivial et qu'il est valable sous des hypothèses plus générales que celles proposées en B.3 ?

Pour un théoricien, l'interprétation qui suit n'est qu'heuristique : elle est évidemment très fortement liée à l'étude de la chaîne incluse (cf. 6.C) associée aux F-transitions définies ci-dessous.

On considère donc les hypothèses données en B.3. On admet qu'il existe un processus X markovien homogène qui évolue à temps continu et qui admet E comme ensemble d'états et a comme matrice d'évolution. Quand le processus X passe de l'état u à l'état v , on suppose que cette transition peut, ou non, être une F-transition ; on suppose que la probabilité pour que cette transition soit (resp. ne soit pas) une F-transition vaut $b_{u,v}/(b_{u,v}+c_{u,v})$ (resp. $c_{u,v}/(b_{u,v}+c_{u,v})$).

$(q_k(t))_{u,v}$ est alors la probabilité, sachant que $X_0 = u$, que $X_t = v$ et qu'il y ait eu exactement k F-transitions entre 0 et t .

La relation (ii) de B.1 (cas où $k = 0$) se démontre comme les équations de Chapman-Kolmogorov (cf. 1.D.2) ; on conseille vivement au lecteur de "redémontrer" cette relation sur un ou deux exemples.

La relation (iii) de B.1 est plus facile à interpréter sous la forme 8B2 : cette relation 8B2 peut s'écrire :

pour tout couple (u,v) d'éléments de E , on a :

$$[q_{k+1}(t)]_{u,v} = \int_0^t \sum_w [q_0(s)]_{u,w} \sum_{w'} m_{w,w'}(q_k)_{w',v}(t-s) ds$$

ce qui est la somme sur w et w' (éléments de E) et l'intégrale sur t de la probabilité qu'entre 0 et s le processus passe de l'état u à l'état w et qu'il n'y ait pas de F-transition, puis qu'il y ait une F-transition qui fait passer de w à

w' entre s et $(s+ds)$ et, enfin, qu'entre s et t il y ait exactement k F-transitions et que le processus X passe de l'état w' à l'état v .

B.5 Remarque

Formellement, ce cours n'utilise presque pas la "théorie générale des processus" (cf. chapitre 10) et repose sur des démonstrations d'algèbre ou d'analyse qui sont à la fois plus simples et plus générales que celles liées à la théorie des processus. Pour autant, l'intérêt de l'interprétation probabiliste ne doit pas être sous-estimé. **D'une part elle est évidemment indispensable au niveau des applications. D'autre part, elle permet de "découvrir" des propriétés qui dépassent le cadre probabiliste (comme le théorème B.1).**

Par exemple, il est évident intuitivement que, dans la preuve heuristique proposée en B.4, on peut intervertir les rôles de q_0 et q_k : $q_{k+1}(t)$ est associé à la probabilité qu'il y ait exactement k F-transitions entre 0 et s puis une F-transition entre s et $(s+ds)$ et, enfin, aucune F-transition entre s et t . On a donc

$$q_{k+1}(t) = \int_0^t q_k(s) m q_0(t-s) ds$$

Evidemment, on peut vérifier directement cette propriété analytique. En fait, on démontre, en raisonnant par récurrence sur j , que, pour $0 \leq j \leq k$, on a :

$$q_{k+1}(t) = \int_0^t q_j(s) m q_{k-j}(t-s) ds$$

ce qui est "heuristiquement évident".

B.6 Cas triangulaire "par blocs"

Soit r une matrice carrée qui admet la décomposition "par blocs" (cf. 5.A.1) suivante :

$$r = \begin{pmatrix} x & o \\ y & z \end{pmatrix}$$

On a alors (cf. [Pel-1]) :

$$e^{rt} = \begin{pmatrix} u & o \\ v & w \end{pmatrix} \text{ avec } u := e^{xt}, w := e^{zt} \text{ et}$$

$$v := \int_0^t e^{z(t-s)} y e^{xs} ds$$

Ceci peut être considéré comme un cas particulier du théorème B.1 avec $m := \begin{pmatrix} o & o \\ y & o \end{pmatrix}$ et $n := \begin{pmatrix} x & o \\ o & z \end{pmatrix}$; on a $e^{nt} = \begin{pmatrix} u & o \\ o & w \end{pmatrix}$, $q_1(t) = \begin{pmatrix} o & o \\ v & o \end{pmatrix}$ et, pour $k > 1$, $q_k(t) = 0$.

C. Flots poissonniens

C.1 Proposition

Soit a , b , c et d quatre matrices indexées par $(E \times E)$ et q une matrice uniligne indexée par $(1 \times E)$. On suppose que $a = b + c$, que d est une matrice diagonale, que $q a = q d$ et que q est un vecteur propre de c c'est à dire qu'il existe un réel u tel que $q c = u q$. On a alors

$$q e^{(b-d)t} = q e^{-ut}$$

Vérification

Aux détails de présentation près, le coeur de cette proposition repose sur la relation bien connue qu'il y a entre les vecteurs propres et les exponentielles de matrice : historiquement, c'est d'ailleurs l'étude des systèmes différentiels linéaires qui a conduit à mettre en évidence l'importance des notions de valeurs et vecteurs propres.

On a ici :

$$q(b-d) = q(a-d) - q c = -u q$$

c'est à dire que q est un vecteur propre de $(b-d)$. On a alors (étude classique)

$$q e^{(b-d)t} = \sum_{k=0}^{\infty} q(b-d)^k t^k / k!$$

mais on vérifie immédiatement par récurrence croissante sur k que $q(b-d)^k = (-u)^k q$ d'où

$$q e^{(b-d)t} = \sum_{k=0}^{\infty} q(-u)^k t^k / k! = q e^{-ut}$$

C.2 Cadre probabiliste

On reprend les hypothèses et conventions introduites en B.3 tout en gardant l'hypothèse très forte $q_c = u q$ introduite en C.1.

Quitte à renormaliser, on peut supposer que $\sum_{v \in E} q_v = 1$: q est alors la probabilité stationnaire associée au processus X (puisque $q(a-d) = 0$) (cf. 2.D.2). $(q e^{(b-d)t})_v$ est alors (cf. B.4) la probabilité d'être dans l'état v à l'instant t et qu'il n'y ait pas eu de F -transition entre 0 et t sachant que q était la loi initiale. **Or, dire que l'on choisit q comme loi initiale, c'est dire que l'on observe le processus en régime stationnaire.** De plus,

$$\sum_{v \in E} (q e^{(b-d)t})_v = e^{-ut} \sum_{v \in E} q_v = e^{-ut}$$

e^{-ut} est donc la probabilité, en régime stationnaire, qu'il n'y ait pas eu de F -transition entre 0 et t . On constate donc, par le calcul, que, en régime stationnaire, les événements A et B sont indépendants avec

$A :=$ il n'y a pas eu de F -transition entre 0 et t
 $B :=$ le processus est dans l'état v à l'instant t

(puisque $\text{proba}(A \cap B) = \text{proba}(A) \text{proba}(B)$).

De plus, en régime stationnaire, le délai avant la première F -transition à partir d'un instant choisi "au hasard" suit la loi exponentielle de paramètre u .

Par ailleurs, posons, pour tout élément v de E :

$$\begin{aligned} p_v &:= \int_0^\infty (q e^{(b-d)t})_v dt \\ &= \sum_{w \in E} \int_0^\infty (q e^{(b-d)t})_w c_{w,v} dt \end{aligned}$$

En raisonnant comme en B.4, cette quantité p_v peut être interprétée de la façon suivante :

p_v est la somme sur w (élément de E) et l'intégrale sur t de la probabilité d'être dans l'état w à t et de ne pas avoir eu de F -transition entre 0 et t et d'avoir une F -transition entre t et $(t+dt)$ qui fait passer de l'état w à l'état v . Autrement dit, si le délai avant la première F -transition est presque sûrement fini, p_v est la probabilité, en régime stationnaire, d'être dans

l'état v "juste après une F-transition". Ceci peut encore s'énoncer " p_v est la probabilité, quand on choisit un instant au hasard puis la première F-transition suivante, d'être dans l'état v immédiatement après cette F-transition".

Or, on a :

$$p = \int_0^{\infty} q e^{-u t} c dt$$

$$= q c / u = q$$

La probabilité juste après la première F-transition est donc la probabilité stationnaire. On peut donc reprendre les raisonnements qui précèdent mais en prenant comme condition initiale l'état juste après la première F-transition.

Notamment, la probabilité juste après la deuxième F-transition – et donc aussi juste après la k -ième F-transition – est toujours q . De plus, en régime stationnaire, le délai entre deux F-transitions est indépendant de l'état immédiatement après la deuxième F-transition et ce délai suit la loi exponentielle de paramètre u .

La suite des F-transitions constitue donc un processus poissonnien (cf. 1.A.3).

C.3 Balance locale

La notion de balance locale a été introduite en 4.B.2. Pour alléger la présentation, considérons le cas particulier où, **pour une classe donnée G de clients il y a balance locale à la station S** c'est à dire que, pour chaque état v , le "taux de probabilité de quitter l'état v avec départ d'un client de classe G de la station S " est égal au "taux de probabilité d'atteindre l'état v avec arrivée d'un client de classe G à la station S ".

Pour tout couple (v, w) d'états, on suppose que $c_{v, w}$ (resp. $b_{v, w}$) est associé aux transitions de l'état v à l'état w pour lesquelles il y a une (resp. il n'y a pas d') arrivée de client de classe G à la station S .

Par ailleurs, on suppose que le taux de service pour la classe G à la station S est constamment égal à u (ce qui implique, notamment, que la station S ne peut jamais être vide).

On a alors $q c = u q$ si q est la probabilité stationnaire : toutes les propriétés indiquées en C.2 sont donc satisfaites. **Notamment, en régime stationnaire, la suite des arrivées de clients de classe G à la station S constitue un processus de Poisson de paramètre u .**

Evidemment, on a la même propriété au niveau des départs puisque le taux de service à la station S est constamment égal à u .

C.4 Remarques

1°) Les résultats de ce paragraphe C s'étendent facilement au cas où l'ensemble E des états est infini dénombrable si on a des familles sommables (cf. 1.D.4) et si le délai entre deux F -transitions est presque sûrement fini (cf. [L2B] et [L2P]).

2°) La mise en évidence de processus poissonniens pour certaines transitions a fait l'objet d'une littérature surabondante : cf., par exemple, [Bur], [Kel], [Bre], [LPS], etc... Plusieurs auteurs ont pensé qu'il serait possible d'approfondir la propriété de balance locale par ce biais.

En fait, autant que l'auteur ait pu en juger, tous les articles dans cette voie - sauf, évidemment, ce qui est associé à une "démonstration" fausse - reposent sur la propriété $q c = u c$ (avec u réel). Notons de plus, que les propriétés démontrées aux paragraphes C.2 et C.3 qui précèdent sont beaucoup plus précises que la seule existence de processus poissonniens sous-jacents. L'auteur pense donc qu'il est illusoire d'espérer découvrir une relation entre la balance locale et l'existence de flux poissonniens en sortie et en entrée autre que ce qui a été noté en C.2 et C.3. Pour d'autres auteurs, le problème, semble-t-il, reste ouvert.

Enfin, comme cela a été noté au chapitre 4, la balance locale peut être considérée en un sens très général et il peut y avoir balance locale à une station et non dans d'autres stations (par exemple, si certaines stations ont des lois de service non exponentielle : cf. 4.F.1).

D. Processus réversibles

D.1 Introduction

L'expression "processus réversible" a deux sens différents dans la littérature récente. Nous prendrons cette expression au sens fort extensivement étudié dans [Kel]. Bien entendu, nous ne chercherons pas à résumer ici tous les résultats exposés dans [Kel]. Conformément à l'esprit général de ce cours, nous allons seulement donner le "cœur" des propriétés d'analyse liées à cette définition sans nous restreindre au cadre probabiliste.

D.2 Pseudo-commutativité

Proposition : Soit b, c, d et q quatre matrices indexées par $(E \times E)$ telles que d et q soient diagonales et $b q = q c$ ("pseudo-commutativité"). On a alors

$$e^{(b-d)t} q = q e^{(c-d)t}$$

Vérification :

q et d étant diagonales on a $q d = d q$ donc $(b-d) q = q (c-d)$. On en déduit, en raisonnant par récurrence croissante sur k , que $(b-d)^k q = q (c-d)^k$. On a donc :

$$\begin{aligned} e^{(b-d)t} q &= \sum_{k=0}^{\infty} (b-d)^k q t^k / k! \\ &= \sum_{k=0}^{\infty} q (c-d)^k t^k / k! = q e^{(c-d)t} \end{aligned}$$

D.3 Processus réversible

Soit X un processus markovien homogène qui évolue à temps continu et qui admet a comme matrice d'évolution (avec $a_{u,u} = 0$ pour tout élément u de E). On dit que X est réversible s'il existe une fonction p définie sur E telle que, pour tout couple (u,v) d'éléments de E ,

$$(8D1) \quad p_u a_{u,v} = p_v a_{v,u}$$

A une constante multiplicative près, p est la **probabilité stationnaire** puisque l'égalité ci-dessus implique

$$\sum_{v \in E} p_u a_{u,v} = \sum_{v \in E} p_v a_{v,u}$$

D.4 Exemples

La propriété de réversibilité est une propriété très forte qui n'est satisfaite que pour quelques réseaux de base.

La file $M/M/1$ est un réseau réversible. Le serveur central comme défini en 4.A.3 est un réseau réversible. Un réseau de Jackson (cf. 4.B.2) est un réseau réversible si la famille $(x_i)_{i \in H}$ introduite en 4.B.1 est telle que, pour tout couple (i,j) d'éléments de H on ait :

$$x_i r_{i,j} = x_j r_{j,i}$$

Il en va de même pour les réseaux BCMP introduits en 4.C.

D.5 Renversement du temps

Proposition :

1°) Soit X un processus réversible. Pour tout couple (u,v) d'états on a, en régime stationnaire :

$$\text{Proba}[X_0 = u \text{ et } X_t = v] = \text{Proba}[X_0 = v \text{ et } X_t = u]$$

2°) De plus, soit F une partie de $(E \times E)$ et G la partie de $(E \times E)$ associée à F comme suit :

(u,v) appartient à G si et seulement si (v,u) appartient à F .

On suppose qu'il n'existe pas d'état u tel que (u,u) appartienne à F .

On dira qu'une transition de u à v est une F -transition (resp. une G -transition) si (u,v) appartient à F (resp. à G). On a alors, pour tout couple (u,v) d'états et en régime stationnaire :

$$\begin{aligned} &\text{Proba}[X_0 = u, X_t = v \text{ et pas de } F\text{-transition entre } 0 \text{ et } t] \\ &= \text{Proba}[X_0 = v, X_t = u \text{ et pas de } G\text{-transition entre } 0 \text{ et } t] \end{aligned}$$

En langage imagé, la loi d'évolution du processus est, "en un certain sens", invariante quand on "retourne le temps".

Vérification :

1°) Soit a la matrice d'évolution associée à X , d la matrice diagonale associée et p la probabilité stationnaire. Soit q la matrice diagonale indexée par $(E \times E)$ et définie par :

$$q_{u,v} := p_u \delta_{u,v}$$

Soit \overline{a} la matrice (indexée par $E \times E$) et définie par :

$$\overline{a}_{u,v} := a_{v,u}$$

\overline{a} est la matrice d'évolution quand on retourne le temps.

La relation 8D1 implique $q a = \overline{a} q$. On peut donc appliquer la proposition D.2 avec $c = a$ et $b = \overline{a}$ ce qui donne

$$q e^{(a-d)t} = e^{(\overline{a}-d)t} q$$

c'est à dire que, pour tout couple (u,v) d'états, on a :

$$p_u (e^{(a-d)t})_{u,v} = (e^{(\overline{a}-d)t})_{u,v} p_v = (e^{(a-d)t})_{v,u} p_v$$

En effet, puisque $(\overline{a-d})$ est la matrice transposée de la matrice $(a-d)$, $e^{(\overline{a-d})t}$ est la matrice transposée de la matrice $e^{(a-d)t}$, d'où la deuxième égalité ci-dessus.

Ceci implique le 1°) (formule de Bayes et propriété de Markov).

2°) On pose :

$$\begin{aligned} c_{u,v} &:= a_{u,v} [1 - 1_F(u,v)] & c'_{u,v} &:= a_{u,v} [1 - 1_G(u,v)] \\ b_{u,v} &:= a_{v,u} [1 - 1_F(u,v)] \end{aligned}$$

autrement dit b est la transposée de c' .

La preuve est alors tout à fait analogue à celle effectuée au 1°). On a $q c = b$ et la proposition D.2 donne :

$$q e^{(c-d)t} = e^{(b-d)t} q$$

c'est à dire que pour tout couple (u,v) d'états :

$$p_u (e^{(c-d)t})_{u,v} = (e^{(b-d)t})_{u,v} p_v = (e^{(c'-d)t})_{v,u} p_v$$

ce qui donne, en régime stationnaire (puisque X est un processus markovien) :

$$\begin{aligned} &\text{Proba } \{X_0 = u \text{ et } X \text{ passe de l'état } u \text{ à l'état } v \text{ entre } 0 \text{ et } t \text{ et il n'y a pas} \\ &\quad \text{de } F\text{-transition entre } 0 \text{ et } t\} \\ &= \text{Proba } \{X_0 = v \text{ et } X \text{ passe de l'état } v \text{ à l'état } u \text{ entre } 0 \text{ et } t \text{ et il n'y a} \\ &\quad \text{pas de } G\text{-transition entre } 0 \text{ et } t\} \end{aligned}$$

ce qui prouve le 2°).

E. Processus de renouvellement

E.1 Définition

Certains processus markoviens particuliers apparaissent dans des études diverses ; on a donc donné un nom spécifique à chacun de ces processus : **processus de renouvellement**, **processus semi-markovien**, **processus régénératif**, etc... Nous ne résumerons pas les propriétés essentielles de ces divers processus : dans ce domaine, la meilleure référence est encore, le chapitre 10 de [Cin] (cf. aussi [Doo]). Nous allons simplement "situer" ce type d'études par rapport aux méthodes proposées

dans ce cours à partir de l'exemple des processus de renouvellement.

Un processus X est un processus de renouvellement (cf. [Cin]) si l'ensemble des états E est fini ou dénombrable et s'il existe une suite croissante $(w(n))_{n>0}$ de "temps d'arrêt" telle que, si on pose $Y_n := X_{w(n)}$ et $s_n := w(n) - w(n-1)$, la propriété suivante soit satisfaite :

pour tout entier n et pour toute séquence $(u_k)_{0 \leq k \leq n+1}$ d'états

$$\begin{aligned} & \text{Proba}[Y_{n+1} = u_{n+1} \text{ et } s_{n+1} \leq t \mid Y_n = u_n] \\ &= \text{Proba}[Y_{n+1} = u_{n+1} \text{ et } s_{n+1} \leq t \mid \forall k \leq n, Y_k = u_k] \end{aligned}$$

E.2 Aspect théorique

Evidemment, du point de vue théorique, ce qui précède ne constitue une définition que si on a d'abord montré la cohérence d'un tel système d'hypothèses, ce qui est loin d'être trivial, notamment si on veut étudier le régime stationnaire. De plus, à chaque fois que l'on introduit de nouvelles hypothèses sur X , il faut vérifier cette cohérence. En fait, du point de vue théorique, le plus simple est de construire un processus de Markov stationnaire naturellement associé à X : pour cela, la première étape est de construire le semi-groupe de transition (cf. le chapitre 1). Vues sous cet angle, les "démonstrations" proposées dans [Cin] abordent le problème "à l'envers" : elles doivent être considérées comme des "preuves heuristiques" au sens donné à cette expression dans ce cours.

E.3 Simulation

Supposons que l'on veuille étudier un processus de renouvellement par simulation : à chaque instant, il faut mettre en mémoire l'état du processus (ce qui impose de choisir E fini) et le délai passé depuis le "dernier" temps d'arrêt u_n (ce qui impose de discrétiser le temps). Comme on l'a déjà noté au chapitre 3, il est souvent préférable de simuler le délai s_n entre deux temps d'arrêt successifs à l'aide d'états fictifs plutôt que de mettre en mémoire une valeur discrétisée de s_n . De plus, quelle que soit la méthode utilisée (états fictifs ou discrétisation du temps), on se ramène à simuler un processus markovien dont l'ensemble des états est fini (ce qui est le cadre général de ce cours).

Avant d'utiliser l'arsenal, nécessairement assez lourd, associé aux processus de renouvellement (ou aux autres processus analogues), il faut donc se demander si cette technique est plus efficace que celle consistant directement à étudier un processus de Markov dont l'ensemble des états est fini.

E.4 Aspect technique

Si le but essentiel de l'étude effectuée est d'obtenir des valeurs moyennes en régime stationnaire, les techniques liées au processus de renouvellement sont en général moins efficaces que les techniques matricielles directes (cf. le chapitre 6 notamment).

Par ailleurs, en régime transitoire, les méthodes "élémentaires" proposées dans les paragraphes précédents doivent être envisagées avant les méthodes plus sophistiquées.

F. Transitoire stationnaire et fiabilité

F.1 Introduction

On a déjà noté (cf. 1.C.5, 1.C.6, 2°) de 2.C.2, 3.D.2, 6.C) qu'on a souvent lieu d'étudier le "**comportement moyen**" d'un régime apparemment transitoire. Ceci est notamment le cas s'il y a des points absorbants (cas d'un jeu à durée finie, étude de fiabilité, etc...).

Du point de vue mathématique, en général, il n'est alors pas nécessaire de calculer explicitement une exponentielle de matrice puisqu'on peut se ramener à une étude en régime stationnaire (ce qui permet d'utiliser les techniques introduites dans les chapitres 5, 6 et 7).

Par exemple, s'il s'agit d'un jeu, dès que le jeu est terminé, on "recommence" avec la probabilité initiale imposée. On se ramène ainsi à l'étude, en régime stationnaire, d'un processus markovien homogène (il n'y a plus d'état absorbant). Evidemment, la probabilité stationnaire de chaque "ex-état absorbant" u est, à une constante multiplicative près, la probabilité que la partie s'achève dans l'état u .

Plus généralement, pour la plupart des applications, il n'est pas nécessaire de calculer explicitement la matrice e^{nt} : nous allons en donner deux exemples.

F.2 Cadre probabiliste

On reprend les hypothèses et conventions introduites en B.3 et on suppose que la matrice c est associée aux transitions de u à v telles que v est un état absorbant (évidemment, on ne suppose pas que la matrice a est irréductible).

Soit z la matrice uni-ligne associée à la loi initiale. La probabilité pour que l'état absorbant w soit atteint entre t et $(t+h)$ vaut

$$\sum_v (z e^{nt})_v c_{v,w} h + \varepsilon(t,h) h$$

avec $\lim_{h \downarrow 0} \varepsilon(t,h) = 0$.

La probabilité pour que le phénomène étudié s'achève dans l'état absorbant w vaut donc :

$$\int_0^\infty (z e^{nt} c)_w dt$$

On suppose

– d'une part que $n = b-d$ est une matrice inversible ; soit r telle que $r n = i$.

– d'autre part que $\lim_{t \rightarrow \infty} e^{nt} = 0$

Dans ce cas, $r e^{nt}$ est la primitive de e^{nt} puisque $\frac{d}{dt}(r e^{nt}) = r n e^{nt} = e^{nt}$.

On a donc

$$\int_0^\infty z e^{nt} c dt = [z r e^{nt} c]_0^\infty = -z r c$$

Rappelons que le calcul de r , qui est la matrice inverse de $b-d$, a été abordé en 6.A.4 et 6.C.2.

F.3 Durée de vie

Dans les études de fiabilité, un point essentiel est, en général, de déterminer le **délai moyen avant d'atteindre un état de panne définitive**.

Considérons les mêmes hypothèses qu'en F.2 ci-dessus. Si l'état absorbant w est atteint entre t et $(t+h)$ (avec h "petit"), la durée de vie vaut t . La durée de vie moyenne avec w comme état absorbant vaut donc

$$\int_0^\infty t (z e^{nt} c)_w dt$$

Or, $(t r e^{nt} - r^2 e^{nt})$ est une primitive de $t e^{nt}$ (ceci se vérifie par dérivation) donc l'intégrale ci-dessus vaut $(z r^2 c)_w$.

Le délai moyen vaut donc

$$\sum_{w \in E} (z r^2 c)_w$$

Evidemment, on étudie de même les temps de séjour (moyens), temps d'attente, etc...

F.4 Chaîne incluse

L'étude qui suit est "heuristique" ; elle complète le paragraphe 6.C.

Dans ce qui précède, on a choisi un "décor" associé aux études de fiabilité compte tenu de l'importance de ce problème pour les utilisateurs. Evidemment, les techniques proposées peuvent être adoptées dans beaucoup d'autres situations.

Reprenons les hypothèses et notations introduites en B.4 et supposons que l'on veuille étudier le processus X par l'intermédiaire de la chaîne incluse associée aux F -transitions. Supposons aussi, comme en F.2, qu'il existe une matrice r telle que $r(b-d) = i$ et que $\lim_{t \rightarrow \infty} e^{nt} = 0$.

D'une part, si p est la probabilité stationnaire pour la chaîne incluse aux instants qui suivent immédiatement une F -transition, la fin de F.2 implique que

$$(8F1) \quad p = p(d-b)^{-1} c$$

D'autre part, $(e^{nt})_{u,v}$ est la probabilité, sachant que $X_0 = u$, d'avoir $X_t = v$ et qu'il n'y ait pas eu de F -transition entre 0 et t .

$\int_0^{+\infty} (e^{nt})_{u,v} dt$ est donc la "place dans le temps" (cf. 6.C.3) occupée par l'état

v quand on observe le processus entre deux F -transitions et si ce processus est dans l'état u à l'instant initial de l'observation.

Puisque p est la probabilité stationnaire pour la chaîne incluse, la "place totale dans le temps" occupée par l'état v , c'est à dire la probabilité stationnaire d'être dans l'état v , vaut, à une constante multiplicative près :

$$\int_0^{+\infty} (p e^{nt})_v dt = (p(d-b)^{-1})_v$$

On retrouve ainsi (aux changements de notations près) le théorème 6.A.2.

On peut aussi expliquer ce résultat de la façon suivante : on pose :

$$q := p(d-b)^{-1}$$

La relation 8F1 implique

$$p(d-b)^{-1}(d-b) = p(d-b)^{-1} c \quad \text{soit} \\ q(d-b) = qc$$

et q est, à une constante multiplicative près, la probabilité stationnaire (si celle-ci est unique).

F.5 Exemple de base

1°) Soit u et v deux réels positifs. Soit X et Y deux processus de Poisson (cf. 1.A.3), nuls en zéro, de paramètres respectifs u et v . Soit w le premier instant (w est un temps d'arrêt : cf. 10.C.7) où l'on a $Y_t = 1$; pour $t < w$ on pose $Z_t := X_t$; pour $t > w$, on pose $Z_t = X_w$. On se propose d'étudier la loi de Z_t puis la limite de cette loi quand t tend vers l'infini.

2°) On pose :

$$p_k(t) := \text{proba} [X_t = k \text{ et } Y_t = 0]$$

$$q_k(t) := \text{proba} [Z_t = k]$$

Les équations de Chapman-Kolmogorov (cf. 1.D.2) s'écrivent :

$$p'_{k+1} = -(u+v) p_{k+1} + u p_k$$

$$q'_k = v p_k$$

Il est facile de calculer la solution de ce système différentiel linéaire en raisonnant par récurrence croissante sur k ce qui donne

$$p_k(t) = e^{-(u+v)t} (ut)^k / k!$$

$$q_k(t) = \int_0^t v p_k(s) ds$$

$$q_0(t) = v [1 - e^{-(u+v)t}] / (u+v)$$

On a aussi (en intégrant par parties) pour $k > 0$:

$$q_k(t) = u q_{k-1}(t) / (u+v) - [v/(u+v)] e^{-(u+v)t} (ut)^k / k!$$

3°) On en déduit :

$$\lim_{t \rightarrow \infty} q_k(t) = v u^k / (u+v)^{k+1}$$

F.6 Un cas tridiagonal par blocs

On reprend les hypothèses considérées en B.4 et F.4. De plus, on suppose qu'il existe une application f définie sur E et à valeurs dans \mathbb{N} : f associe à chaque état le nombre de clients "dans une station". On suppose que (cas tridiagonal par blocs) $|f(e) - f(e')| > 1$ implique $a(e, e') = 0$.

On suppose de plus que c est associé "aux départs" des clients précités c'est à dire que $a(e, e') = c(e, e')$ si et seulement si $f(e') = f(e) - 1$. On note $q = p(d-b)^{-1}$ la probabilité stationnaire à une constante multiplicative près.

Enfin, et ceci est une hypothèse très restrictive, on suppose qu'il existe une constante positive h telle que $f(e') = f(e) + 1$ implique $a(e, e') = h$ (h est le "taux d'arrivée") : cette hypothèse est satisfaite pour la file M/PH/1.

Pour tout k entier positif, soit $E(k)$ l'ensemble des éléments e de E tels que $f(e) = k$. On a alors (cf. [Lem]... et 5.E.3, 2°) $x = y$ avec :

$$x := \sum_{e \in E(k)} \sum_{e' \in E(k+1)} q(e) a(e, e')$$

$$y := \sum_{e \in E(k+1)} \sum_{e' \in E(k)} q(e) a(e, e')$$

Or,

$$x = h \sum_{e \in E(k)} q(e)$$

et

$$y := \sum_{e \in E(k)} (qc)_e = \sum_{e \in E(k)} p(e)$$

La probabilité marginale stationnaire d'avoir k clients dans la station est égale (puisque proportionnelle) à la probabilité marginale stationnaire de la chaîne incluse associée aux instants qui suivent immédiatement un départ.

F.7 File unique

Soit E l'ensemble des entiers k avec $0 \leq k \leq n$ (avec $n > 1$, n fixé). Soit u et v deux fonctions définies sur E et positives. On suppose que $v(0) := u(n) = 0$ et que, dans tous les autres cas, $v(k)$ et $u(k)$ sont strictement

positifs. Soit a la matrice indexée par $(E \times E)$ et définie par :
 si $0 < k \leq n$, $a(k, k-1) := v(k)$ et $a(k-1, k) := u(k-1)$ et, dans tous les autres cas, $a(j, k) := 0$.

Les exemples proposés en 1.E.4, 9.C.1 et 9.E.1 sont des cas particuliers du cadre considéré dans ce paragraphe F.7. Soit d la matrice diagonale associée à a (comme en 2.D.3) c'est à dire que, quel que soit k , $0 \leq k \leq n$, $d(k, k) := u(k) + v(k)$.

Soit c la matrice indexée par $(E \times E)$ dont tous les termes sont nuls sauf $c(1, 0) := v(1)$ et soit $b := a - c$. La matrice $\exp. [(b-d) t]$ n'a pas de forme analytique simple ; même dans le cas particulier des files M/M/1 ou M/M/8, l'écriture de cette exponentielle n'est pas triviale.

Soit q la matrice inverse de $(d-b)$ (avec les notations considérées en F.2, $q = -r$). **Le point remarquable dans cet exemple est que q a une forme analytique très simple.** Plus précisément, quel que soit i élément de E , on construit $q(i, j)$ par récurrence croissante sur j par $q(i, 0) := 0$,

pour $1 \leq j \leq i$, $q_{i,j} = (1 + q_{i,j-1} u_{j-1}) / v_j$

pour $j > i$, $q_{i,j} = q_{i,j-1} u_{j-1} / v_j$

Notamment, $q_{i,1} = 1 / v_1$.

Notons d'ailleurs que ce résultat s'étend sans modification formelle au cas d'une file infinie : la condition d'ergodicité s'écrit alors

$$\sum_{j=1}^{\infty} q(i, j) < +\infty$$

(cette condition ne dépend évidemment pas de i compte tenu des formules ci-dessus).

En appliquant la formule donnée en F.3, il est alors facile de calculer le délai moyen de retour en 0 à partir d'un état j quelconque.

F.8 Remarques

1°) Il est évidemment beaucoup plus facile de calculer r , la matrice inverse de $(b-d)$, que la matrice e^{nt} . En fait, la plupart des calculs qui interviennent en fiabilité ont pour but de déterminer des "valeurs moyennes" : comme ci-dessus, **l'exponentielle e^{nt} est alors à utiliser comme intermédiaire mathématique mais n'a pas à être calculée explicitement.**

2°) Dans ce chapitre, nous nous sommes limités au cas où l'ensemble des états est fini par souci de clarté et de simplicité. Comme on l'a déjà noté en C.4, ces résultats peuvent s'étendre au cas où E est infini dénombrable.

Chapitre 9

Fonction génératrice**A. Notions de base****A.1 Introduction**

Cet "outil mathématique", sans support intuitif, que l'on appelle la **fonction génératrice** - on dit aussi la fonction génératrice de Laplace ou la transformée en z - est l'analogue, pour les variables aléatoires entières (les variables de "comptage"), de la transformée de Fourier pour les variables aléatoires admettant une densité de probabilité ; en fait, pour le théoricien, ceci est beaucoup plus qu'une analogie.

Or, les problèmes abordés dans ce cours font très souvent intervenir des **variables de "comptage"** tel le nombre de clients dans une station. Il n'est donc pas surprenant que cette technique de la fonction génératrice soit à la base de centaines d'articles permettant de "calculer analytiquement" des quantités caractéristiques associées à des processus markoviens discrets.

Il est évidemment impossible de reproduire tous ces résultats dans ce chapitre. De plus, en ce qui concerne les applications, une formule analytique n'est intéressante que si elle conduit à des calculs exploitables et/ou à des algorithmes implémentables sur ordinateur.

L'objet de ce chapitre est donc essentiellement de donner **les idées et les techniques de base à partir d'exemples où la fonction génératrice est encore un outil opérationnel**. Notons toutefois que, malgré la multiplicité des articles sur ce thème, l'utilisation du développement au voisinage de $z=1$ (cf. A.3, B.6, C.5, D.4 et E.4) et l'utilisation de e^{ut} quand u est une matrice (cf. D.3 et F.3) sont des techniques qui ont été peu ou pas exploitées auparavant.

Pour la commodité du lecteur, nous allons d'abord résumer les notions de base liées à la fonction génératrice : pour plus de détails cf. [Pel-6], [Met-1], etc...

A.2 Cas à une dimension

1°) On dit que X est une **variable aléatoire entière** si l'ensemble des valeurs de X est contenu dans l'ensemble des entiers positifs. La loi (de probabilité) d'une telle variable aléatoire X est caractérisée par la suite $(q_n)_{n \geq 0}$ avec $q_n := \text{Proba}[X = n]$.

2°) **La fonction génératrice associée à X - plus précisément à la loi de X - est alors définie formellement par :**

$$(9A1) \quad G_X(z) = \sum_{n=0}^{\infty} q_n z^n$$

3°) Soit D le domaine (c'est à dire l'ensemble des complexes z) où la série entière ci-dessus est convergente. Puisque $\sum_{n=0}^{\infty} q_n = 1$, ce domaine D contient le disque de rayon unité et on a $G_X(1) = 1$.

4°) Quand elle est finie, on vérifie que l'espérance de X est la valeur au point $z = 1$ de la dérivée de $G_X(z)$ par rapport à z : si cette espérance est infinie, cette dérivée $G'_X(z)$ tend en croissant vers l'infini quand z réel tend en croissant vers 1. On a un résultat analogue pour la variance de X.

5°) **La fonction génératrice caractérise la loi de X** en ce sens que, si deux variables aléatoires X et X' ont la même fonction génératrice, alors X et X' ont la même loi de probabilité.

6°) Comme pour la transformée de Fourier, la propriété la plus importante de la fonction génératrice a trait **aux sommes de variables aléatoires indépendantes**. Plus précisément, si X et Y sont deux variables aléatoires entières indépendantes et si on pose $S := X+Y$, on a :

$$G_S(z) = G_X(z) G_Y(z)$$

Cette propriété s'étend au cas de n variables, $n > 2$.

7°) Evidemment, la fonction génératrice est particulièrement utile si c'est une fonction analytique "simple" ce qui est le cas pour la loi binomiale et la loi de Poisson. Plus précisément,

a) $G_X(z) = (1+d(z-1))^n$ si et seulement si

$$\text{Proba } [X = k] = C_n^k d^k (1-d)^{n-k} \text{ pour } 0 \leq k \leq n$$

b) $G_Y(z) = e^{a(z-1)}$ si et seulement si

$$\text{Proba } [Y = k] = e^{-a} a^k / k!$$

8°) Soit X et X' deux variables aléatoires entières de lois respectives q et q' . Soit a et a' deux nombres positifs tels que $a+a' = 1$. Soit Y la variable aléatoire dont la loi est définie par :

$$\text{Proba } [Y = n] = a q_n + a' q'_n$$

On a alors $G_Y = a G_X + a' G_{X'}$.

Ce résultat s'étend au cas de n variables, $n > 2$.

On peut d'ailleurs prouver très simplement le 6°) en utilisant cette propriété.

A.3 Développement en $z = 1$

Dans ce paragraphe, on suppose que le rayon de convergence de la série associée à la fonction génératrice $G(z)$ est strictement supérieur à 1. Dans ce cas, la théorie des fonctions analytiques (cf. [Car]) permet d'affirmer que **$G(z)$ est développable en série entière au voisinage de $z = 1$** .

Autrement dit, il existe une suite de réels $(r_k)_{k \geq 0}$ telle que la série de terme général $r_k(z-1)^k$ a un rayon de convergence u strictement positif et

$$G_X(z) = \sum_{k \geq 0} r_k(z-1)^k \quad \text{pour } |z-1| < u$$

On a évidemment $r_0 = G_X(1) = 1$; de plus, pour $|z-1| < u$, on peut dériver indéfiniment termes à termes la série ci-dessus ; notamment, $r_k / k!$ est la valeur de la dérivée k -ième de G (par rapport à z) prise au point $z = 1$. Ceci implique que tous les coefficients r_k sont positifs. De plus (cf., par exemple, [Pel-1]), $r_1 = E(X)$ et $\text{var}(X) = 2 r_2 + r_1 - (r_1)^2$. Rappelons enfin que la fonction génératrice des variables aléatoires entières usuelles (cf., par exemple, A.2, 7°)) est souvent une fonction "analytiquement simple" de $(z-1)$ et donc tous les coefficients r_k sont connus explicitement.

Pour alléger la présentation, on se limite, dans ce cours, au cas où $G(z)$ est développable en série entière au voisinage de $z = 1$, ce qui suffit pour la quasi-totalité des applications : si on veut effectuer une étude mathématique plus fine, on étudie $G(z)$ pour z réel, $0 \leq z \leq 1$ et on fait tendre z vers 1 en croissant.

A.4 Cas multidimensionnel

Soit k' un entier fixé, $k' > 1$; soit K l'ensemble des entiers k tels que $1 \leq k \leq k'$. Pour alléger les notations on pose $E := \mathbb{N}^K$. Soit X une variable aléatoire à valeurs dans E dont la loi est définie par la famille $(q(n))_{n \in E}$ avec $q(n) = \text{Proba}[X = n]$.

Pour tout élément $z := (z(k))_{k \in K}$ de \mathbb{C}^K , la valeur en z de la fonction génératrice de X est définie par :

$$G_X(z) = \sum_{k \in K} \sum_{n(k)=0}^{\infty} q_n \prod_{k \in K} z(k)^{n(k)}$$

avec, évidemment, $n := (n(1), \dots, n(k) \dots n(k'))$.

Toutes les propriétés de la fonction génératrice dans le cas unidimensionnel s'étendent au cas multidimensionnel.

A.5 Constante de normalisation

On a vu en 4.B.4 que la fonction génératrice permet de démontrer facilement une propriété algébrique des sommes de produits "homogènes" : ceci donne un algorithme performant pour calculer, notamment, la constante de normalisation quand la solution a une forme produit. Il s'agit là d'une utilisation spécifique de la fonction génératrice dont nous ne reparlerons pas dans ce chapitre.

B File M/PH/1

B.1 Introduction

Dans ce paragraphe B, on étudie la "**chaîne incluse**" associée à la file M/PH/1. D'une part, dans ce cas, la fonction génératrice satisfait à la relation 9B3 qui est exceptionnellement simple. D'autre part, ce paragraphe donne un éclairage complémentaire de l'étude de base, beaucoup plus générale, proposée en 6.C (voir, aussi, 5.E). Enfin ce paragraphe montre, dans un cadre élémentaire, quels sont les problèmes théoriques qui se posent : il peut donc être considéré comme une **introduction au chapitre 10 qui suit**.

B.2 Hypothèses mathématiques (cf. 1.F.4)

1°) Soit J un ensemble fini (l'ensemble des "états fictifs") et w un élément fixé de J (état pivot). L'ensemble E des états comprend d'une part l'élément $(w, 0)$ (cas où la file est vide), d'autre part tous les couples (j, k) avec j élément de J et k entier strictement positif (k est associée à la longueur de la file).

L'état de la file est modélisé par un processus X markovien homogène qui admet E comme ensemble d'états et qui évolue à temps continu. Soit a la matrice d'évolution (cf. 1.D.1), c'est à dire que :

$$a(e, e') := \lim_{h \downarrow 0} \frac{1}{h} \text{Proba} [X_{t+h} = e' \mid X_t = e]$$

Soit f (resp. g) une fonction positive définie sur $(J \times J)$ (resp. sur J) et v un réel positif. On définit a de la façon suivante :

$$a((w,0),(w,1)) := v$$

et pour $k > 0$:

$$a((j,k),(j',k)) := f(j,j')$$

$$a((j,k),(w,k-1)) := g(j)$$

$$a((j,k),(j,k+1)) := v$$

Enfin, $a(e,e')$ est nul dans tous les autres cas.

2°) Avec cette modélisation, v est le taux d'arrivée, $f(j,j')$ est le taux de passage de l'état fictif j à l'état fictif j' et $g(j)$ est le taux de service quand on est dans l'état fictif j , une fin de service provoquant un retour de l'état fictif à l'état pivot w .

Rappelons que toute modélisation de la loi de service ne comportant qu'un ensemble fini d'états fictifs peut se ramener à une modélisation avec pivot. Rappelons aussi que n'importe quelle loi peut être approchée, avec autant de précision qu'on le souhaite, par une loi admettant une modélisation par états fictifs, l'ensemble de ces états fictifs étant fini.

3°) En fait, les fonctions f et g ne seront pas utilisées explicitement par la suite. Le but de la modélisation proposée au 1°) est, d'une part, de montrer comment l'étude de ce paragraphe est un cas particulier des études antérieures (cf. 6.B et 6.C notamment) et, d'autre part, de donner un cadre d'hypothèses non ambigu.

B.3 Chaîne incluse

Soit $(u(n))_{n \geq 0}$ la suite des "instants" de départs : ces instants sont des "temps d'arrêt" (cf. 10.C.7). Pour ce paragraphe B.3, une formalisation mathématique sophistiquée (comme cela est proposé au chapitre 10) n'est pas absolument indispensable.

Soit Z_n la variable aléatoire associée à la longueur de la file à l'instant $u(n)$: cette taille de la file caractérise l'état du processus X (à l'instant $u(n)$) puisque l'état fictif est l'état pivot. Soit A_n la variable aléatoire associée au nombre d'arrivées entre l'instant $u(n)$ et l'instant $u(n+1)$. On a évidemment $Z_{n+1} = Z_n + A_n - 1$.

Soit B_n la variable aléatoire associée au nombre d'arrivées pendant la

durée d'un service ; si $Z_n = 0$, la loi de A_n est la loi de $1+B_n$; si $Z_n > 0$, la loi de A_n est la loi de B_n , indépendamment de la valeur de Z_n (sachant $Z_n > 0$).

Soit H la fonction génératrice associée à la variable aléatoire B_n ; ce qui précède implique (cf. A.2, 8°) :

$$(9B1) \quad G_{Z(n+1)} = \text{Proba}[Z_n = 0] \cdot H + \sum_{k=1}^{\infty} \text{Proba}[Z_n = k] z^{k-1} H$$

ce qui peut aussi s'écrire, en posant $r_n := \text{Proba}[Z_n = 0]$:

$$(9B2) \quad G_{Z(n+1)} = r_n H + (G_{Z(n)} - r_n) H/z$$

B.4 Régime stationnaire

1°) Considérons alors le cas, de loin le plus important au niveau des applications, où on observe X en régime stationnaire. Intuitivement, il est "évident" que la propriété S est satisfaite avec :

$S :=$ "la loi de $Z(n+1)$ est la même loi que la loi de $Z(n)$ ".

Du point de vue de l'utilisateur, il n'y a pas d'inconvénient à "admettre" cette propriété S .

2°) Par contre, attention, si une étude "**théorique**" mathématique est basée sur cette propriété S , elle ne peut être considérée comme **complète** que si la **cohérence** de cette propriété S et des autres hypothèses du problème a été établie. Or, prouver cette cohérence fait nécessairement appel à un arsenal théorique de haut niveau (ce qui n'était pas le cas au paragraphe B.3).

Par exemple, pour que la propriété S soit satisfaite, il faut que, pour tout n , $u(n)$ soit "presque sûrement fini" mais cela ne suffit pas.

3°) Admettons donc la propriété S et soit G la fonction génératrice associée à $Z(n)$. La relation 9B2 devient (avec $r = r_n$) :

$$G = r H + (G - r) H/z \quad \text{soit}$$

$$(9B3) \quad G = r H(1-z) / (H-z)$$

4°) On a vu en 8.F.6 que la loi de $Z := Z(n)$ est la loi stationnaire de la longueur de la file.

B.5 Calcul de H

1°) Considérons d'abord le cas où la durée de service B suit une loi exponentielle de paramètre u . On a vu en 8.F.5 que, dans ce cas, la probabilité d'avoir k arrivées durant un service est :

$$\text{Proba}[B = k] = u v^k / (u+v)^{k+1}$$

La fonction génératrice associée est donc :

$$\begin{aligned} H &= u / (u+v-vz) \text{ ce qui implique} \\ H-z &= (u-vz)(1-z) / (u+v-vz) \text{ et} \\ G &= ru / (u-vz) \end{aligned}$$

Or, pour $z = 1$, $G(1) = 1$ donc :

$$\begin{aligned} r &= (u-v) / u \text{ et} \\ G &= (u-v) / (u-vz) \text{ soit} \end{aligned}$$

$$\text{Proba}[Z_n = k] = [1-v/u](v/u)^k$$

et on note que (cf. 2.C.2, 3°) et 8.F.6)

$$\text{Proba}[Z_n = k] = \text{Proba}[X_t = k]$$

2°) Considérons maintenant le cas général pour lequel il existe une

famille $(y_i)_{i \in I}$ de réels positifs telle que $\sum_{i \in I} y_i = 1$ et une famille $(Y_{i,k})$, $i \in I$,

$k \in K(i)$ de variables aléatoires indépendantes suivant chacune la loi exponentielle de paramètre $u(i,k)$ et telles que la loi de la durée de service Y est la même que la loi de

$$\sum_{i,k} y_i Y_{i,k}$$

Autrement dit, la loi de Y est la combinaison convexe (de coefficients $(y_i)_{i \in I}$) des "produits de convolution" $\sum_k Y_{i,k}$.

Avec la probabilité y_i , le nombre d'arrivées durant un service sera la somme sur k du nombre d'arrivées durant le délai associé à la variable aléatoire $Y_{i,k}$: or, on a vu au 1°) que la fonction génératrice $H_{i,k}$ de ce nombre d'arrivées vaut

$$H_{i,k} = u_{i,k} / (u_{i,k} + v - vz)$$

Le nombre d'arrivées durant un service est la combinaison convexe (de coefficients $(y_i)_{i \in I}$) de la somme (le produit de convolution) sur k du nombre d'arrivées durant le délai associé à $Y_{i,k}$. La fonction génératrice H

associée à ce nombre d'arrivées vaut donc :

$$H = \sum_i y_i \prod_{k \in K(i)} H_{i,k}$$

B.6 Calcul des moments

Considérons le cas où le rayon de convergence de la série associée à la fonction génératrice H est strictement supérieur à 1 (cf. A.3). On pose

$$F := (H-1) / (z-1)$$

F est développable en série entière au voisinage de $z = 1$ et les valeurs en $z = 1$ des dérivées de F (par rapport à z) sont positives (et, en général, "connues").

La relation 9B3 devient :

$$(9B4) \quad G = r H / (1-F) = r H / [1-F(1) - (F-F(1))]$$

ou encore (puisque $G(1) = 1$) :

$$(9B5) \quad G = H \sum_{k=0}^{\infty} [F-F(1)]^k / [1-F(1)]^k$$

et la série ci-dessus converge au voisinage de $z = 1$ (cf. [Car]) : évidemment, $F(1) = E(B) < 1$ (sinon il y a "explosion" et la propriété S énoncée en B.4 n'est pas satisfaite).

La relation 9B4 (qu'on peut aussi utiliser sous la forme 9B5) permet de calculer le développement en série entière de $(z-1)$ pour la fonction G en fonction de ce même développement pour la fonction H (et donc la fonction F).

A titre illustratif, appliquons cette méthode pour calculer la moyenne et la variance de Z .

On suppose que

$$\begin{aligned} G &= 1 + \sum_{k \geq 1} p_k (z-1)^k \quad \text{et} \\ H &= 1 + \sum_{k \geq 1} q_k (z-1)^k \\ \text{donc } F &= \sum_{k \geq 0} q_{k+1} (z-1)^k \end{aligned}$$

La relation 9B5 donne, par identification et pour $q_1 < 1$:

$$(9B6) \quad p_1 = q_1 + q_2 / (1-q_1)$$

$$(9B7) \quad p_2 = q_2 + (q_1 q_2 + q_3) / (1 - q_1) + q_2^2 / (1 - q_1)^2$$

La relation 9B6 peut aussi s'écrire :

$$(9B8) \quad E(Z) = \frac{1}{2} \{ E(B) + \text{var}(B) / [1 - E(B)] \}$$

Cette relation est appelée la formule de Pollaczek-Khintchine (cf. B.4.4°) et 8.F.6).

Rappelons de plus que $\text{var}(Z) = 2 p_2 + p_1 - (p_1)^2$.

Par ailleurs, dans les cas considérés en B.5, il est "facile" de calculer $E(B)$ et $\text{var}(B)$. Enfin, tous les calculs qui précèdent pourraient être implémentés sur ordinateur : par exemple, il serait "facile" de mettre au point un algorithme qui donne "formellement" la famille $(p_k)_{k \geq 0}$ (jusqu'à un certain rang) en fonction de la famille $(q_k)_{k \geq 0}$.

C. Processus de naissances et de morts

C.1 Hypothèses

1°) Nous allons maintenant étudier **le régime transitoire du processus de naissances et de morts au sens strict**. D'une part ce processus est l'un des plus connus et des plus anciens de toute cette théorie. D'autre part son étude va permettre d'introduire à partir d'un exemple simple les idées et les techniques cruciales dans toute la suite de ce chapitre.

2°) Le modèle mathématique est le suivant : l'ensemble E des états est l'ensemble des entiers naturels et on étudie un processus X markovien homogène qui admet E comme ensemble d'états et qui évolue à temps continu à partir de l'instant $t = 0$.

Soit a la matrice d'évolution (cf. 1.D.1) associée à ce processus, c'est à dire que :

$$a(e, e') := \lim_{h \downarrow 0} \frac{1}{h} \text{Proba} [X_{t+h} = e' \mid X_t = e]$$

Soit u et v deux réels strictement positifs. On pose, quel que soit $k \geq 0$: $a(k, k+1) := u$, $a(k+1, k) := v$ et $a(e, e') := 0$ dans tous les autres cas.

On pose $p_k(t) := \text{Proba} [X_t = k]$.

3°) L'interprétation est alors la suivante : X est associé à une population (par exemple une famille de microbes) où il y a des naissances et des morts ; u (resp. v) est le taux de naissance (resp. de mortalité).

C.2 Equations de Chapman-Kolmogorov

Puisqu'il s'agit d'un exemple de base et pour la commodité du lecteur, rappelons comment on obtient ces équations (cf. 1.D.2). Pour $h > 0$ et $t > 0$, on a :

$$\text{Proba } [X_{t+h} = k] = \sum_{i=1}^4 x_i$$

où x_1 est la probabilité que $X_t = k$ et qu'il n'y ait pas de changement d'état entre t et $t+h$, x_2 est la probabilité que $X_t = k+1$ et que l'état passe de $k+1$ à k entre t et $t+h$, x_3 est la probabilité que $X_t = k-1$ (qui est nulle si $k = 0$) et que l'état passe de $k-1$ à k entre t et $t+h$ et, enfin, x_4 est la probabilité qu'il y ait au moins deux changements d'états entre t et $t+h$. On a :

$$x_1 = p_k(t) [1 - (u+v) k h] + h \varepsilon_1(h)$$

$$x_2 = (p_{k+1}(t) v(k+1) h + h \varepsilon_2(h)$$

$$x_3 = p_{k-1}(t) u(k-1) h + h \varepsilon_3(h)$$

$$x_4 = h \varepsilon_4(h)$$

avec, quel que soit i , $\lim_{h \downarrow 0} \varepsilon_i(h) = 0$.

On en déduit :

$$[p_k(t+h) - p_k(t)] / h = -(u+v) k p_k(t)$$

$$+ v(k+1) p_{k+1}(t) + u(k-1) p_{k-1}(t) + \varepsilon(h)$$

avec $\lim_{h \downarrow 0} \varepsilon(h) = 0$ et la convention $p_{-1}(t) := 0$.

En faisant tendre h vers zéro, on obtient les équations de Chapman-Kolmogorov :

$$(9C1) \quad \frac{d}{dt} p_k(t) = -(u+v) k p_k(t) + v(k+1) p_{k+1}(t) + u(k-1) p_{k-1}(t)$$

C.3 Fonction génératrice

La relation 9C1 peut aussi s'écrire :

$$(9C2) \quad z^k \frac{d}{dt} p_k(t) = -(u+v) z y_1 + v y_2 + u z^2 y_3 \quad \text{avec}$$

$$y_1 = k z^{k-1} p_k(t) = \frac{\partial}{\partial z} [z^k p_k(t)]$$

$$y_2 = (k+1) z^k p_{k+1}(t) = \frac{\partial}{\partial z} [z^{k+1} p_{k+1}(t)]$$

$$y_3 = (k-1) z^{k-2} p_{k-1}(t) = \frac{\partial}{\partial z} [z^{k-1} p_{k-1}(t)]$$

On pose $G(z,t) := \sum_{k=0}^{\infty} z^k p_k(t)$.

En additionnant les relations 9C2 pour k variant de 0 à l'infini, on obtient, pour $|z| < 1$:

$$(9C3) \quad \frac{\partial}{\partial t} G(z,t) = [-(u+v)z + v + uz^2] \frac{\partial}{\partial z} G(z,t)$$

C.4 Propriétés générales de la solution

1°) La relation 9C3 est appelée "équation aux dérivées partielles linéaires du premier ordre". L'étude de telles équations fait l'objet d'une littérature abondante (cf., par exemple, [Val]). Les propriétés qui suivent ont un caractère général relativement à de telles équations.

2°) D'une part, on vérifie immédiatement que toute combinaison linéaire de solutions de 9C3 est aussi une solution de 9C3.

3°) D'autre part, dans la relation 9C3, la fonction G n'apparaît que par l'intermédiaire de ses dérivées partielles : on vérifie alors immédiatement que, si G_1 et G_2 sont deux solutions de 9C3, $G = G_1 G_2$ est aussi une solution de 9C3.

4°) **Cette propriété a une interprétation probabiliste précise :** considérons deux "populations" disjointes et les processus X_1 et X_2 associés à ces populations ; on suppose que ces deux processus sont indépendants l'un de l'autre et satisfont aux propriétés données précédemment (les valeurs de u et v étant les mêmes pour ces deux processus). On vérifie alors immédiatement que le processus $X := X_1 + X_2$ satisfait aussi aux propriétés données précédemment. En d'autres termes, les "descendants" associés aux deux populations initiales $X_1(0)$ et $X_2(0)$ évoluent indépendamment les uns des autres.

5°) Notamment, soit G_1 (resp. G_n) la fonction génératrice associée au processus X^1 (resp. X^n) tel que $\text{Proba}[X^1(0) = 1] = 1$ (resp. $\text{Proba}[X^n(0) = n] = 1$) : on a alors : $G_n = (G_1)^n$.

La fonction génératrice G du processus X peut donc s'écrire :

$$(9C4) \quad G(z,t) = \sum_{k=0}^{\infty} \text{Proba}[X_0 = k] \cdot [G_1(z,t)]^k$$

Il suffit alors de déterminer G_1 (cf, par exemple, [Bar]).

6°) On peut aussi utiliser A.3 si la fonction génératrice associée à la condition initiale a un rayon de convergence strictement supérieur à 1. Dans ce cas, soit $H(z,t)$ la solution particulière de 9C3 telle que $H(z,0) = z-1$. La solution cherchée de 9C3 est alors

$$(9C5) \quad G(z,t) = 1 + \sum_{k \geq 1} r_k [H(z,t)]^k$$

la famille $(r_k)_{k \geq 1}$ étant déterminée par (cf. A.3) :

$$G(z,0) = 1 + \sum_{k \geq 1} r_k (z-1)^k$$

7°) Si on s'intéresse essentiellement aux probabilités $\text{Proba}[X_t = k]$, il vaut mieux utiliser le 5°). Par contre, si on s'intéresse essentiellement aux divers moments de X_t , il vaut mieux procéder comme au 6°). Par exemple, 9C5 implique :

$$\frac{\partial G}{\partial z} = \sum_{k \geq 1} k r_k [H(z,t)]^{k-1} \frac{\partial H}{\partial z}$$

mais, en général, $H(1,t) = 0$, ce qui implique :

$$\frac{\partial}{\partial z} G(1,t) = r_1 \frac{\partial}{\partial z} H(1,t) \quad \text{soit}$$

$$\frac{\partial}{\partial z} G(1,t) = \frac{\partial}{\partial z} G(1,0) \frac{\partial}{\partial z} H(1,t) \quad \text{ou encore}$$

$$E(X_t) = E(X_0) \frac{\partial}{\partial z} H(1,t)$$

8°) Evidemment, tout ce qui précède s'étend aux exemples analogues.

C.5 Solution associée à $z-1$

1°) On se propose de déterminer la fonction $H(z,t)$ telle que $H(z,0) = z-1$ et (cf. 9C3) :

$$(9C6) \quad \frac{\partial H}{\partial t} = [-(u+v)z + v + uz^2] \frac{\partial H}{\partial z}$$

2°) On utilise alors la "méthode des caractéristiques". Quand l'équation à étudier est de la forme

$$\frac{\partial H}{\partial t} = Q \frac{\partial H}{\partial z}$$

cette méthode consiste à commencer par résoudre le système différentiel

$$(9C7) \quad dt = -dz/Q$$

Dans 9C6, $Q = -(u+v)z + v + uz^2 = (z-1)(uz - v)$ et la résolution de 9C7 peut s'écrire :

$$(u-v) dt = -\frac{dz}{z-1} + \frac{u dz}{uz-v} \quad \text{soit}$$

$$K e^{(u-v)t} = (uz-v)/(z-1) \quad \text{soit}$$

$$(9C8) \quad K = e^{(v-u)t} (uz-v)/(z-1)$$

où K est une "constante arbitraire".

3°) La "méthode des caractéristiques" consiste à dire que la solution générale H de 9C6 est une "fonction arbitraire" de la fonction K définie par 9C8, soit

$$(9C9) \quad H(z,t) = f[e^{(v-u)t}(uz-v)/(z-1)]$$

Ici, on cherche la solution de 9C6 telle que $H(z,0) = z-1$ et 9C9 implique :

$$\begin{aligned} z-1 &= f[(uz-v)/(z-1)] \\ &= f([u(z-1)+u-v]/[z-1]) \end{aligned}$$

On inverse donc la fonction f, considérée comme fonction de $z-1$, suivant la méthode classique, c'est à dire que l'on calcule y en fonction de x sachant que :

$$x = (uy+u-v)/y$$

ce qui équivaut à :

$$(9C10) \quad y = (u-v)/(x-u)$$

Autrement dit, on obtient la fonction H en remplaçant, dans 9C10, x par K (cf. 9C8) soit

$$(9C11) \quad H(z,t) = (u-v) / \{ [e^{(v-u)t}(uz-v)/(z-1)] - u \}$$

D'une part, $H(z,t)$ est une solution de 9C6 puisque c'est une fonction de K ; d'autre part, pour $t = 0$, $H(z,0) = z-1$ (le calcul qui précède est "fait pour cela"). Il suffit de reporter 9C11 dans 9C5 pour connaître G.

La relation 9C11 implique $H(1,t) = 0$ (quelque soit t). Il est alors facile de calculer la valeur en $z = 1$ de $\frac{\partial H}{\partial z}$ ce qui donne, avec $e := e^{(v-u)t}$:

$$(u-v) e(u-v) / e^2(u-v)^2 \text{ soit } 1/e .$$

On en déduit (cf. C.4, 7°) :

$$E(X_t) = E(X_0) e^{(u-v)t}$$

ce qui est un résultat classique que l'on obtient ainsi très rapidement quelle que soit la condition initiale.

D. Cas presque linéaire en z

D.1 Hypothèses

Soit h' un entier avec $h' \geq 2$. Soit H l'ensemble des entiers j tels que $1 \leq j \leq h'$. Soit s et s' deux fonctions positives définies sur H et r une fonction positive définie sur $(H \times H)$. On suppose que, pour tout élément j de H , $r_{j,j} = 0$ et $\sum_k r_{j,k} = 1$.

L'ensemble E des états est une partie de \mathbb{N}^H . Etant donné $n := (n_j)_{j \in H}$ élément de \mathbb{N}^H , j et k éléments de H avec $n_k > 0$, on définit la notation $(n+e_j-e_k)$ comme suit (cf. 4.A.2) :

$n+e_j-e_k$ est l'élément n' de \mathbb{N}^H tel que $n'_j := 1+n_j$, $n'_k := n_k-1$ et, pour $i \neq j$ et $i \neq k$, $n'_i := n_i$.

On définit $n-e_k$ de façon analogue.

On suppose que pour tout élément n de E et pour tout couple (j,k) d'éléments de H (resp. tout élément k de H) avec $n_k s_k r_{k,j} > 0$ (resp. $n_k s'_k > 0$), l'élément $(n+e_j-e_k)$ (resp. $(n-e_k)$) appartient aussi à E .

La matrice d'évolution (cf. 1.D.1) est définie par :

$$a(n, (n+e_j-e_k)) := s_k n_k r_{k,j}$$

$$a(n+e_j, n) := s'_j(1+n_j)$$

$a(n, n') = 0$ s'il n'existe pas de couple (j,k) , $j \neq k$, tel que $n' = n+e_j-e_k$ et s'il n'existe pas d'indice i tel que $n' = n-e_i$.

Ce modèle peut être interprété exactement comme en 4.B (s'_i étant

associé à un départ vers l'extérieur) : par contre les hypothèses sont beaucoup plus restrictives qu'en 4.B puisqu'on suppose ici que le **taux de service dans chaque station est proportionnel au nombre de clients dans la station**. Notons toutefois que, puisqu'on s'intéresse au régime transitoire (et non pas au régime stationnaire), il n'y a pas lieu de supposer que la matrice r est irréductible (il peut y avoir des stations "absorbantes").

Soit $p_t(n)$ la probabilité pour le processus d'être dans l'état n , n appartenant à E , à l'instant t . Si n n'appartient pas à E , on pose $p_t(n) := 0$.

D.2 Equations de Chapman-Kolmogorov

On considère ces équations à l'instant t ; pour alléger les notations le paramètre t est omis. Ces équations (cf. 1.D.2) s'écrivent pour l'état

$n := (n_j)_{j \in H}$:

$$(9D1) \quad \frac{d}{dt} p(n) = - \sum_i (s_i + s'_i) n_i p(n) + \sum_{i,j} s_j (1+n_j) r_{j,i} p(n+e_j-e_i) \\ + \sum_i s'_i (1+n_i) p(n+e_i)$$

Pour tout couple (n, z) avec n élément de \mathbb{N}^H et z élément de \mathbb{C}^H , on pose :

$$f(z, n) := p(n) \prod_{h \in H} z_h^{n(h)}$$

Après multiplication par $\prod_{h \in H} z_h^{n(h)}$, la relation 9D1 peut s'écrire :

$$(9D2) \quad \frac{\partial}{\partial t} f(z, n) = - \sum_i (s_i + s'_i) z_i \frac{\partial}{\partial z(i)} [f(z, n)] \\ + \sum_{i,j} s_j z_i r_{j,i} \frac{\partial}{\partial z(j)} [f(z, n+e_j-e_i)] \\ + \sum_i s'_i \frac{\partial}{\partial z(i)} f(z, n+e_i)$$

On introduit alors la fonction génératrice :

$$G(z) := \sum_{n \in E} f(z, n)$$

En sommant sur n élément de E , les relations 9D2 impliquent (commencer par le cas où $\text{card}(H) = 2$ pour s'en convaincre) :

$$(9D3) \quad \frac{\partial}{\partial t} G(z) = \sum_i [s'_i - (s_i + s'_i) z_i] \frac{\partial G(z)}{\partial z(i)} + \sum_{i,j} s_j z_i r_{j,i} \frac{\partial G(z)}{\partial z(j)}$$

On considère $z := (z_1, \dots, z_{h'})$ comme une matrice uni-ligne ; on désigne - avec une notation abusive mais évocatrice - par $\frac{\partial G}{\partial z}$ la matrice uni-colonne ($h' \times 1$) dont la i -ième ligne vaut $\frac{\partial G}{\partial z(i)}$; enfin soit u (resp. v) la matrice ($h' \times h'$) (resp. ($1 \times h'$)) définie par :

$$u_{i,i} := -(s_i + s'_i), \quad v_{1,i} := s'_i$$

et, si $i \neq j$ $u_{i,j} := s_j$ $r_{j,i}$.

Les relations 9D3 peuvent alors s'écrire matriciellement comme suit :

$$(9D4) \quad \frac{\partial G}{\partial t} = (z u + v) \frac{\partial G}{\partial z}$$

D.3 Résolution

On suppose, soit que $v = 0$, soit que u est inversible. On a évidemment les mêmes propriétés qu'en C.4. Par ailleurs, soit w une matrice uni-colonne "fixe" (indépendante de z et t). Si on pose :

$$W = z e^{ut} w + v u^{-1} e^{ut} w - v u^{-1} w, \quad \text{on a :}$$

$$\frac{\partial W}{\partial t} = z u e^{ut} w + v e^{ut} w \quad \text{et}$$

$$\frac{\partial W}{\partial z} = e^{ut} w \quad (\text{puisque } \frac{\partial z}{\partial z} \text{ est la matrice unité})$$

ce qui implique :

$$\frac{\partial W}{\partial t} = z u \frac{\partial W}{\partial z} + v \frac{\partial W}{\partial z}$$

et W est une solution de 9D4.

Soit w_i la matrice uni-colonne ($h' \times 1$) dont tous les termes sont nuls sauf le i -ième qui vaut 1 et soit $W_i := (z + v u^{-1}) e^{ut} w_i - v u^{-1} w_i$, c'est à dire que W_i est la i -ième colonne de la matrice $(z + v u^{-1}) e^{ut} - v u^{-1}$.

La solution de 9D4 associée au problème étudié est alors :

$$(9D5) \quad G = \sum_{n \in E} \text{Proba}[X_0 = n] \prod_{h=1}^{h'} (W_h)^{n(h)}$$

D.4 Calcul des moments

La relation 9D5 est à utiliser quand on veut calculer les probabilités

des divers états à un instant donné. Par contre, si on veut étudier les divers "moments" (et donc les moyennes, variances, etc...), il est préférable de considérer $G(z,t)$ au voisinage de $z_h = 1$.

Plus précisément, appelons $\overline{0}$ (resp. $\overline{1}$) l'élément n de \mathbb{N}^H tel que, quelque soit h élément de H , $n_h := 0$ (resp. 1) ; appelons E' l'ensemble \mathbb{N}^H privé du point $\overline{0}$. Supposons que la fonction génératrice associée à l'état initial (pour $t = 0$) soit développable en série entière au voisinage de $z = \overline{1}$ (avec un rayon de convergence strictement positif). Soit $(q(n))_{n \in E'}$ la famille de réels positifs définie par :

$$G(z,0) = 1 + \sum_{n \in E'} q(n) \prod_{h \in H} (z_h - 1)^{n(h)}$$

On a alors :

$$G(z,t) = 1 + \sum_{n \in E'} q(n) \prod_{h \in H} (W_h - 1)^{n(h)}$$

où W_h est défini comme en D.3.

D.5 Exemples

1°) Notons d'abord que, dans ce paragraphe D, il n'y a pas de "naissance" (au sens du paragraphe C) car la résolution serait alors beaucoup plus compliquée techniquement.

2°) Les exemples les plus classiques sont issus de la biologie. Par exemple, on considère une population qui comprend plusieurs classes d'individus - bien portants, contaminés par un virus, guéris - et, dans chaque classe, il peut y avoir plusieurs groupes (race, facteur rhésus, âge, etc...). Si on peut supposer que chaque individu évolue (éventuellement) **"indépendamment des autres"**, les hypothèses de ce paragraphe D sont satisfaites. Notons que si les tailles de ces populations sont importantes, les hypothèses ci-dessus sont à peu près satisfaites même si la loi de passage d'un état à un autre (pour chaque individu) n'est pas exactement une loi exponentielle.

3°) Evidemment, on a exactement les mêmes propriétés mathématiques si on considère un parc de machines, chaque machine pouvant prendre plusieurs états - bon état, légèrement dégradée, en panne grave, réparée, etc... - sous réserve que ces **diverses machines évoluent "indépendamment" les unes des autres**.

4°) Enfin, nous verrons dans l'exemple qui suit que la résolution d'une équation aux dérivées partielles linéaire générale utilise souvent la

résolution de l'équation "sans second membre" associée, c'est à dire une équation où G n'intervient que par l'intermédiaire de ses dérivées partielles : on peut alors se ramener à la situation étudiée dans ce paragraphe D si l'équation générale considérée est "presque linéaire en z ".

E File M/M/ ∞

E.1 Hypothèses

Soit u et v deux réels strictement positifs. L'ensemble E des états est l'ensemble des entiers. La matrice d'évolution du processus X (cf. 1.D.1) est définie par, quel que soit $k \geq 0$:

$$\begin{aligned} a(k, k+1) &= u \\ a(k+1, k) &= v(k+1) \end{aligned}$$

Dans tous les autres cas, $a(e, e') := 0$.

On pose $p_{-1}(t) := 0$ et, pour $k \geq 0$, $p_k(t) := \text{Proba}[X_t = k]$.

Les équations de Chapman-Kolmogorov s'écrivent :

$$p'_k = -(u + v k) p_k + v(k+1) p_{k+1} + u p_{k-1}$$

On pose :

$$G(z, t) := \sum_{k \geq 0} p_k(t) z^k$$

Les équations ci-dessus impliquent :

$$(9E1) \quad \frac{\partial}{\partial t} G = v(-z+1) \frac{\partial G}{\partial z} + u(z-1) G$$

E.2 Résolution générale

L'équation 9E1 est encore une "équation aux dérivées partielles linéaire" mais, dans cette équation, G apparaît "**explicitement**" (et non pas seulement par l'intermédiaire de ses dérivées partielles).

La résolution générale d'un tel type d'équation peut se faire en utilisant la "méthode des caractéristiques" (cf., par exemple, [Val]). On peut aussi (ce qui est équivalent) procéder comme suit.

On vérifie d'abord qu'une combinaison linéaire de solutions est une solution ; un produit de deux solutions n'est pas, en général, une solution. Par contre, soit G une solution de 9E1 et H une solution de l'équation "sans second membre" associée, c'est à dire que :

$$(9E2) \quad \frac{\partial H}{\partial t} = -v(z-1) \frac{\partial H}{\partial z}$$

On vérifie alors immédiatement que HG est une solution de 9E1. On dit que la "solution générale" de 9E1 est le produit de la "solution générale" de 9E2 et d'une "solution particulière" de 9E1.

Pour la résolution de 9E2 on procède comme dans les paragraphes précédents ; on considère donc (cf. la relation 9C7 en C.5, 2°)) le système différentiel linéaire

$$dt = dz / v(z-1)$$

et $H = f[e^{-vt}(z-1)]$ est la "solution générale" de 9E2.

E.3 Solution particulière

En général, les équations de Chapman-Kolmogorov conduisent à une équation aux dérivées partielle de la forme :

$$(9E3) \quad \frac{\partial G}{\partial t} = A(z) \frac{\partial G}{\partial z} + B(z) G$$

où les fonctions A et B ne dépendent que de z (et non pas de t). Dans ce cas, on peut choisir la "solution particulière" G de 9E3 sous la forme d'une fonction G de z seulement. Autrement dit on cherche une fonction K de z telle que :

$$(9E4) \quad A(z) \frac{\partial K}{\partial z} + B(z) K = 0$$

L'équation 9E4 est une "équation différentielle linéaire à coefficients variables A(z) et B(z)". En général, c'est à dire si les fonctions A(z) et B(z) sont "compliquées", l'équation 9E4, qu'on peut ramener à une "quadrature", n'a pas de solution explicite simple. Par contre, il existe une très grande variété d'exemples pour lesquels on connaît cette solution.

Dans le cas de l'équation 9E1, $A(z) = v(-z+1)$ et $B(z) = u(z-1)$ et l'équation 9E4 admet la solution

$$K = e^{uz/v}$$

La "solution générale" de 9E1 est donc

$$(9E5) \quad G = e^{uz/v} f[e^{-vt}(z-1)]$$

où f est une fonction "quelconque".

E.4 Calcul final

Il y a plusieurs façons d'expliciter la fonction génératrice G(z,t) associée à X_t .

1°) Pour l'étude des probabilités d'états, on peut procéder comme suit :

Soit $(p_k)_{k \geq 0}$ la famille de réels telle que :

$$(9E6) \quad \sum_{k=0}^{\infty} p_k z^k = e^{-uz/v} G(z,0)$$

On sait que cette famille existe et que la série entière associée à un rayon de convergence au moins égal à 1 : c'est le produit des séries entières associées à $e^{-uz/v}$ et $G(z,0)$ respectivement. De plus, ce produit est souvent une fonction "simple" et la famille p_k est "facile" à déterminer. La solution cherchée G peut alors s'écrire :

$$(9E7) \quad G(z,t) = e^{uz/v} \left\{ \sum_{k=0}^{\infty} p_k [e^{-vt}(z-1) + 1]^k \right\}$$

2°) Pour étudier les moments, il vaut mieux procéder comme suit : on suppose que la série entière associée à $G(z,0)$ a un rayon de convergence strictement supérieur à 1 ; les fonctions $e^{-uz/v}$ et $G(z,0)$ sont alors développables en série entière au voisinage de $z = 1$; soit $q_k z^k$ le terme général de la série produit ; c'est à dire que la famille $(q_k)_{k \geq 0}$ est définie par :

$$(9E8) \quad \sum_{k=0}^{\infty} q_k (z-1)^k = e^{-uz/v} G(z,0)$$

La solution cherchée G peut alors s'écrire :

$$(9E9) \quad G(z,t) = e^{uz/v} \left\{ \sum_{k=0}^{\infty} q_k [e^{-vt}(z-1)]^k \right\}$$

Evidemment, on a à la fois 9E7 et 9E9.

En considérant les valeurs en $z = 1$ de G , $\frac{\partial G}{\partial z}$ et $\frac{\partial^2 G}{\partial z^2}$ on obtient :

$$q_0 = e^{-u/v}$$

$$E(X_t) = (u/v) + e^{u/v} q_1 e^{-vt}$$

$$\frac{\partial^2 G}{\partial z^2}(z=1) = (u/v)^2 + 2(u/v) e^{u/v} q_1 e^{-vt} + 2 e^{u/v} q_2 e^{-2vt}$$

Evidemment, on détermine q_1 et q_2 en considérant les relations ci-dessus en $t = 0$. Par exemple :

$$q_1 = e^{-u/v} [E(X_0) - (u/v)]$$

d'où :

$$E(X_t) = (u/v) + [E(X_0) - (u/v)] e^{-vt}$$

F. Etude "abstraite"

F.1 Préliminaire

1°) Il ne faudrait pas que le lecteur se fasse des illusions sur l'efficacité de la fonction génératrice au vu des exemples qui précèdent. En fait, **il est tout à fait exceptionnel de pouvoir calculer explicitement la fonction génératrice. Or, s'il est impossible de mener l'étude jusqu'à ce calcul explicite de G, une résolution partielle n'a aucun intérêt.**

A titre illustratif, dans ce paragraphe F, nous allons considérer un cas où on peut déterminer G de façon "abstraite" et, pourtant, ce résultat est, en général, peu utile (cf. cependant la remarque à la fin de ce paragraphe F).

2°) L'exemple considéré est une généralisation des paragraphes D et E conjointement.

3°) Notons d'abord qu'au paragraphe D on aurait pu supposer $v = 0$ sans perte de généralité. En effet, puisqu'on peut choisir une matrice r qui n'est pas irréductible, au lieu de supposer qu'il y a des départs vers l'extérieur (ce qui correspond aux coefficients s_i) il est équivalent - en ce qui concerne les applications sinon les équations - d'introduire une station supplémentaire (l'extérieur) pour laquelle le taux de service est nul. Dans ce qui suit, on va donc se limiter au cas $v = 0$.

4°) La généralisation par rapport au paragraphe D va consister à supposer qu'il peut y avoir des arrivées en provenance d'un extérieur autre que celui évoqué au 3°) ci-dessus. Le taux pour de telles arrivées est fixe : il ne s'agit donc pas de "naissances" comme cela a été considéré au paragraphe C.

F.2 Equations

On définit h' , H , s , r , u , E et z comme au paragraphe D. De plus, soit b une fonction positive définie sur H . On considère $b := (b_h)_{h \in H}$ comme une matrice unicolonne ($h' \times 1$) et on pose $b' := \sum_{h \in H} b_h$.

La matrice d'évolution a (cf. 1.D.1) est définie par :

$$a(n, (n+e_j - e_k)) := s_k n_k r_{k,j} \quad (\text{comme au paragraphe D})$$

$$a(n, n+e_j) := b_j$$

$$a(e, e') := 0 \text{ dans tous les autres cas.}$$

La fonction génératrice G est définie comme en D.2 et on montre, facilement que les équations de Chapman-Kolmogorov sont équivalentes à la relation suivante :

$$(9F1) \quad \frac{\partial G}{\partial t} = z u \frac{\partial G}{\partial z} + (zb - b') G$$

F.3 Résolution

On suppose que la matrice u est inversible et diagonalisable c'est à dire qu'il existe une matrice $(h' \times h')$ inversible c telle que $u = c d c^{-1}$ où d est une matrice diagonale inversible : on pose $d(i) := d_{i,i}$ et $y := (y_i)_{i \in H} = zc$. On vérifie immédiatement que, pour tout élément i de H , $y_i^{b'/d(i)}$ est une solution "particulière" de l'équation

$$z u \frac{\partial G}{\partial z} - b' G = 0$$

Notons h_i le terme de $z u^{-1} b$ qui ne dépend que de y_i et ne dépend donc pas de y_j pour $j \neq i$,

et $(y_i)^{b'/d(i)} \exp. [-h_i]$ est une solution particulière de l'équation

$$z u \frac{\partial G}{\partial z} + (zb - b') G = 0$$

La solution générale de 9F1 est alors (cf. §E ci-dessus) une combinaison linéaire des solutions suivantes (de 9F1)

$$[(zc)_i]^{b'/d(i)} \exp. [-h_i] F_i$$

où F_i est une fonction "arbitraire" (cf. D.3) des colonnes de la matrice $z e^{ut}$.

On est donc dans un cas où G est "connue" explicitement en tant que solution "générale" de 9F1. En fait, en considérant des cas particuliers, le lecteur peut constater que le "calcul final" de G associé à une condition initiale donnée (pour $t = 0$) conduit à des résultats presque inutilisables (en général).

Par contre, pour des valeurs très particulières (et donc peu réalistes) des paramètres (par exemple $b' = k d(i)$, k entier) ou des conditions initiales adéquates, G a une forme simple. Plus généralement, **l'intérêt essentiel des techniques proposées dans ce chapitre est de donner des résultats exacts dans des cas particuliers sophistiqués : on peut ensuite utiliser ces résultats pour "tester" la validité de programmes généraux de calculs (exacts ou approchés), d'"hypothèses simplificatrices", etc...**

Chapitre 10

Introduction à la théorie**A. Préliminaires****A.1 Objet du chapitre**

Toutes les démonstrations - au sens strict - données dans les chapitres précédents sont des démonstrations complètes et n'utilisent que quelques résultats classiques et relativement élémentaires d'algèbre ou d'analyse. Cela signifie-t-il que la "théorie des probabilités" est inutile quand on se restreint à étudier un "processus markovien homogène" dont l'ensemble E des états est fini et quand on ne s'intéresse qu'aux résultats actuellement effectivement utilisables pour les applications concrètes ?

Les théorèmes établis dans les chapitres précédents sont, en fait, des théorèmes d'algèbre ou d'analyse élémentaires dont le domaine d'application dépasse le seul cadre probabiliste. Toutefois, pour la plupart d'entre eux, l'interprétation probabiliste est fondamentale, notamment pour établir le lien entre ces théorèmes et une modélisation du réel qui laisse une place "au hasard".

De plus, aux chapitres 2, 3, 6 et 8, ce cours propose des "démonstrations heuristiques" : en général, (cf., notamment, 6.C), ces démonstrations ne sont pas mathématiquement indispensables puisque, par ailleurs, des démonstrations (qui n'utilisent pas l'interprétation probabiliste) complètes et rigoureuses sont données. Toutefois, en ce qui concerne la formule de Little, seule une "preuve heuristique" a été proposée (cf. 2.F.4). De plus, même si une preuve heuristique n'est pas incontournable, il serait tout de même intellectuellement rassurant d'en donner une formulation "rigoureuse".

L'objet de ce chapitre est donc essentiellement de "situer" les travaux théoriques qui permettent de transformer les "preuves heuristiques" évoquées précédemment en résultats mathématiques rigoureux : on se limite au cas des processus markoviens homogènes dont l'ensemble des états est fini.

A.2 Rigueur mathématique

La rigueur mathématique est une notion à la fois relative et absolue. Sur le fond, elle est absolue en ce sens qu'une preuve rigoureuse le restera au cours des siècles futurs, même si sa formulation doit être profondément affinée. Sur la forme, notamment en ce qui concerne la précision de la

conceptualisation, les "preuves", évidemment "rigoureuses" proposées par Pascal (milieu du 17ème siècle) ou même Laplace (début du 19ème siècle) sont considérées aujourd'hui comme insuffisamment formalisées.

De même, en 7.A.2, on a noté que, aujourd'hui, la notion de "définition mathématique" donnait lieu à plusieurs variantes dans son interprétation : ce point sera peut-être amélioré dans l'avenir.

La notion de rigueur est particulièrement cruciale en probabilités : c'est, en effet le seul domaine où l'axiomatique mathématique s'appuie sur des considérations intellectuelles – la conceptualisation du hasard – qui ne relèvent ni des mathématiques proprement dites (comme l'arithmétique) ni des sciences expérimentales (comme la géométrie).

Quand on parle de rigueur en théorie des probabilités, est-ce relativement aux concepts probabilistes "de bon sens" - telle la "formule de Bayes" - où est-ce relativement à une axiomatique considérée mathématiquement comme indiscutable ? Dans ce cours, c'est évidemment cette deuxième signification qui est adoptée. Lorsque les preuves proposées sont mathématiquement incomplètement formulées, on y dit que ce sont des "preuves heuristiques".

A.3 L'axiome du choix

Au début du siècle (environ) les mathématiciens se sont convaincus de la nécessité d'introduire un nouvel "axiome" au niveau du raisonnement mathématique, axiome dont la nécessité est beaucoup moins évidente que celle associée, par exemple, au raisonnement par récurrence classique. Au milieu du siècle, les logiciens ont prouvé que cet axiome est "non décidable" ce qui signifie, en simplifiant, que l'utiliser ne peut pas conduire à une contradiction.

La théorie probabiliste utilise abondamment les grands théorèmes d'analyse classique et donc l'axiome du choix : ce point de départ pourrait être remis en cause dans l'avenir mais ceci ne pourrait que déplacer les problèmes.

A.4 Loi de probabilité

Quand on effectue une construction mathématique associée à un problème probabiliste, une des premières étapes est de construire une probabilité sur un espace adéquat. Par exemple, si on veut étudier la loi équirépartie sur $[0,1]$ (segment unité de l'axe réel), on démontre qu'il existe une probabilité p unique définie sur la tribu des boréliens de $[0,1]$ et telle que $p((a,b)) = b-a$ pour tout intervalle (a,b) tel que $a < b$.

Nous ne rappellerons pas les diverses étapes de telles constructions, ni même les définitions d'une probabilité ou de la tribu des boréliens, que

l'on peut trouver dans tous les ouvrages de base (cf. par exemple, [Met-1]). Par contre, rappelons que c'est l'axiome du choix qui implique qu'il n'existe pas de probabilité p' qui prolonge p et qui soit définie sur toutes les parties de $[0,1]$. Ceci introduit des problèmes techniques importants, qu'on appelle les problèmes de mesurabilité, mais n'a pas d'incidence "concrète" immédiate. Evidemment, un lecteur qui souhaite approfondir ses connaissances en théorie des processus doit, d'abord, acquérir quelques connaissances en théorie de la mesure et de l'intégration.

A.5 Espace fondamental

En général, en théorie des probabilités, si cela est possible, "une" loi de probabilité est construite sur un espace Ω suffisamment riche pour prendre en compte "toutes les réalisations possibles du hasard" eu égard aux problèmes étudiés. Ensuite, on étudie des applications - des variables aléatoires - définies sur cet espace Ω .

Le choix de cette démarche est essentiellement technique : il est plus facile, en général, d'étudier diverses fonctions définies sur un même espace Ω que diverses lois de probabilité définies sur des espaces distincts.

Cette démarche est également adoptée en théorie des processus. Plus précisément, **une étape essentielle et indispensable de l'étude "théorique" d'un phénomène qui évolue "aléatoirement" au cours du temps est la construction de l'espace fondamental associé. Cette construction se fait à l'aide d'un ou plusieurs processus canoniques.**

B. Processus canonique

B.1 Loi d'un processus

Considérons le cas où on étudie un processus $X := (X_t)_{t \in T} : T$ est associé au temps et X_t est la variable aléatoire associée à l'état du processus à l'instant t . On suppose que l'ensemble E des états possibles à chaque instant t est fixe. Rappelons la définition de la loi de X donnée en 1.B.2 : c'est la donnée, pour toute séquence $S := (t(1), \dots, t(n))$ croissante d'éléments de T de la loi de $X_S := (X_{t(1)}, \dots, X_{t(n)})$. On dit aussi que cette loi constitue le "système des probabilités conjointes". **Cette famille de probabilités conjointes possède des propriétés de compatibilité "évidentes"** : si S et S' sont deux parties finies de T (deux séquences) avec S contenue dans S' , la loi de X_S est la loi image de $X_{S'}$ par la projection canonique $\Pi_{S,S'}$ de $E^{S'}$ sur E^S . On dit alors que c'est un système projectif.

B.2 Le théorème de Kolmogorov-Bochner

Le processus X peut être considéré comme une "variable aléatoire" à valeurs dans E^T : il est alors techniquement important de pouvoir définir la loi de X en tant que loi de probabilité sur une tribu de parties de E^T . Soit \mathcal{F} la plus petite tribu de parties de E^T qui rend "mesurables" toutes les applications $X \rightarrow X_t$ (t parcourant T).

Le théorème de Kolmogorov-Bochner donne des conditions suffisantes pour que le système des probabilités conjointes admette un prolongement p qui soit une loi de probabilité sur \mathcal{F} : ce prolongement est appelé la limite projective.

Une version de ce théorème, suffisante pour la plupart des applications est donnée et prouvée au chapitre 7 de [Met-1]. Notamment, un tel prolongement existe toujours si l'ensemble E est fini (et si on a bien un système projectif). Ce prolongement p est évidemment unique. **On dit que (E^T, \mathcal{F}, p) est un processus canonique : $\Omega := E^T$ est l'espace fondamental associé à X .**

Il est ensuite possible de définir et d'étudier d'autres processus Y définis sur cet espace fondamental : l'application qui à chaque élément ω considéré comme élément de Ω associe le "même" élément de E^T considéré comme une fonction définie sur T et à valeurs dans E est appelée la projection canonique. Comme dans le cas des variables aléatoires, et même davantage, cette présentation peut sembler, au début, assez artificielle : en fait, un peu d'expérience convaincra le lecteur novice que c'est la seule voie praticable si on ne veut pas se restreindre aux types d'études proposées dans les chapitres précédents.

B.3 Exemple

Considérons le cas où $T = \mathbb{R}^+$ et X est un processus de Markov : le système des probabilités conjointes est caractérisé par la loi initiale (la loi à l'instant $t = 0$) et le semi-groupe de transition (cf. 1.B.4). Si l'ensemble E des états possibles à chaque instant est fini et si le processus est homogène, ce semi-groupe est de la forme e^{bt} (avec $b := d-a$) : cf. 1.D. Le système des probabilités conjointes est un système projectif (par construction) et on peut appliquer le théorème de Kolmogorov-Bochner.

B.4 Régime stationnaire

Considérons le cas où $T = \mathbb{R}$ et X est un processus markovien homogène qui admet E comme ensemble d'états. Dire qu'on étudie X "en

régime stationnaire" c'est dire que la loi de X_t ne dépend pas de t .

Supposons E fini et soit e^{bt} la matrice qui donne la loi d'évolution entre s et $s+t$. Soit q une matrice uniligne telle que $q e^{bt} = q$, quel que soit t : on sait qu'il existe une et une seule telle matrice, à une constante multiplicative près, si la matrice d'évolution a est irréductible (cf. 2.D). On vérifie immédiatement que q et e^{bt} induisent un système projectif (cf. le 2°) de 2.D.6) qui admet une limite projective (cf. B.2).

La loi p , définie sur (E^T, \mathcal{F}) , associée à cette limite projective est appelée la loi stationnaire du processus X .

Attention : Il ne faut pas se tromper sur l'ordre dans lequel est effectuée la construction qui précède : on détermine le semi-groupe de transition et la matrice q avant de construire le processus stationnaire canonique. L'existence et l'unicité de q ne peuvent donc être prouvées que par des méthodes d'algèbre ou d'analyse (comme cela a été fait aux chapitres 1 et 2 si E est fini) et non après la construction du processus stationnaire canonique.

Dans les ouvrages spécialisés, la partie "construction du processus canonique" est souvent omise car considérée comme "prérequis" : pour autant, il ne faut pas oublier qu'une étude mathématique qui repose sur un ensemble d'hypothèses ne peut être considérée comme rigoureuse que si l'on a démontré (ou, du moins, s'il est possible de démontrer) que ce système d'hypothèses est cohérent : c'est en ce sens qu'on a déjà noté (cf. 8.E.2) que les preuves proposées dans [Cin] sont, pour un théoricien, des preuves "heuristiques".

B.5 Automate exponentiel stationnaire

On pose $T := \mathbb{R}$.

Considérons un automate qui oscille constamment entre les positions 1 et 2, le délai entre deux changements d'état suivant la loi exponentielle de paramètre u . On a vu en 1.E.6 que le semi-groupe de transition est facile à calculer, plus précisément :

$$\text{Proba}[X_t = i \mid X_0 = j] = \frac{1}{2} [1 - e^{-2ut} + 2 \delta_i^j e^{-2ut}]$$

(δ_i^j = symbole de Kronecker classique) .

On vérifie immédiatement que la loi stationnaire (unique) q est $q(1) = q(2) = 1/2$.

Si on pose $E := \{1, 2\}$, cette loi stationnaire q et ce semi-groupe de

transition induisent une limite projective p sur $(E^{\mathbb{R}}, \mathcal{F})$ - où \mathcal{F} est la plus petite tribu qui rend mesurable toutes les "projections" à chaque instant t . On dira que $(E^{\mathbb{R}}, \mathcal{F}, p)$ est le processus canonique associé à l'automate exponentiel stationnaire.

Evidemment on peut restreindre la construction qui précède au cas $T = \mathbb{R}^+$.

B.6 Processus de Poisson

On pose $T := \mathbb{R}^+$.

Considérons un processus croissant, à valeurs dans \mathbb{N} , que l'on observe à partir de l'instant $t = 0$, nul en $t = 0$, dont les seules transitions possibles sont une augmentation d'une unité et tel que le délai entre deux telles transitions suit la loi exponentielle de paramètre u .

On a vu en 1.E.3 que le semi-groupe de transition associé est facile à calculer ; plus précisément :

$$\text{Proba}[X_{s+t} = j+k \mid X_s = j] = (ut)^k e^{-ut} / k!$$

$$\text{La loi initiale est } \text{Proba}[X_0 = j] = \delta_0^j.$$

Si on pose $E := \mathbb{N}$, cette loi initiale et ce semi-groupe de transition induisent une limite projective p sur $(E^{\mathbb{R}^+}, \mathcal{F})$. On dit que $(E^{\mathbb{R}^+}, \mathcal{F}, p)$ est le processus de Poisson canonique issu de 0.

Evidemment, cela n'a pas de sens de parler du processus de Poisson stationnaire : par contre, en régime stationnaire, on a souvent besoin de la "dérivée" du processus de Poisson - en ce sens que seuls sont à prendre en compte les instants de sauts et non la valeur exacte du processus. Ceci peut être formulé en termes de mesures aléatoires ou de processus ponctuels (notamment, cf. [Nev] et [Ch1]). Dans les cas élémentaires, l'automate exponentiel stationnaire construit en B.5 suffit.

C. Processus cadlag et temps d'arrêt

C.1 Introduction

Dans ce paragraphe, on suppose qu'on a construit l'espace fondamental (Ω, \mathcal{F}, p) (qui est donc un espace probabilisé) et que l'on étudie un processus X défini sur cet espace : X est une application définie $(\Omega \times T)$ et à valeurs dans E ; E est l'ensemble des états et T est associé au temps.

X peut être l'application canonique ; plus généralement X est une fonction de l'application canonique.

Pour chaque élément ω de Ω , l'application $t \rightarrow X(\omega, t)$ est appelée **la trajectoire** associée à ω : en général, cette trajectoire ne possède aucune propriété de continuité. Or, tant sur le plan technique - par exemple pour pouvoir utiliser efficacement les temps d'arrêt - que du point de vue conceptuel - les réalisations du hasard, si l'ensemble E des états est fini, sont "fixes par morceaux" - **il est indispensable de se restreindre à des trajectoires ayant certaines propriétés de continuité**. Le but de ce paragraphe est de donner les idées et les références de base associées.

C.2 Modification d'un processus

Soit X et X' deux processus définis sur le même espace fondamental (Ω, \mathcal{F}, p) et qui admettent T comme ensemble associé au temps : X et X' sont donc deux applications définies sur $(\Omega \times T)$. On dit que X' est une **modification** de X si, quel que soit t élément de T , $\text{Proba}[X_t \neq X'_t] = 0$. Ceci implique évidemment que X et X' ont la même loi (le même système de probabilités conjointes). Par conséquent, si X' a de "meilleures" propriétés que X , on peut étudier la loi de X à partir de l'étude de X' .

C.3 Processus indistingables

Soit X et X' deux processus définis sur le même espace fondamental (Ω, \mathcal{F}, p) et qui admettent le même ensemble T associé au temps. On dit que X et X' sont **indistingables** si il existe un élément A de \mathcal{F} tel que $p(A) = 0$ et tel que, quel que soit (ω, t) élément de $(\Omega \setminus A) \times T$, $X(\omega, t) = X'(\omega, t)$.

Si X' est indistinguable de X , X' est une modification de X mais la "réciproque est fausse".

C.4 Processus cadlag

Soit X un processus défini sur l'espace fondamental (Ω, \mathcal{F}, p) , qui admet T comme ensemble associé au temps et E comme espace d'états. On suppose que E est muni d'une topologie (si E est fini on le munit de la topologie discrète).

On dit que X est cadlag si, pour tout élément ω de Ω , l'application (la trajectoire) $t \rightarrow X(\omega, t)$ est continue à droite et admet une limite à gauche en tout point t de T . Si E est fini, cela signifie que chaque trajectoire de X est continue à droite et "fixe par morceaux". On dit que X est cadlag à l'indistingabilité près si X est indistinguable d'un processus cadlag.

C.5 Existence de modification cadlag

La notion de processus cadlag est importante d'une part parce qu'elle suffit de point de vue technique, d'autre part parce que pour **tous les processus classiques X** - notamment tous ceux que l'on a étudiés précédemment - **il y a une modification Y de X qui est un processus cadlag.**

L'idée essentielle pour prouver l'existence d'une modification cadlag est due à Doob dans le cas des martingales. On peut trouver une version à la fois simple et très générale de ce théorème d'existence au paragraphe 1.16 de [MeP]. Notons que ce théorème n'utilise pas la "théorie des martingales". Notons aussi que le théorème 4 du chapitre 7 de [Met-1] correspond à une formulation trop restrictive pour beaucoup d'applications (la théorie a beaucoup évolué depuis la publication de [Met-1]).

C.6 Filtration

On suppose que T , l'ensemble associé au temps, est un sous-ensemble de \mathbb{R} (T est donc totalement ordonné). Soit (Ω, \mathcal{F}, p) un espace fondamental. On appelle **filtration** une famille croissante $(\mathcal{F}_t)_{t \in T}$ de sous-tribus de \mathcal{F} .

Pour tout élément t de T , \mathcal{F}_t peut être interprétée comme la famille des événements qui ne font pas intervenir le futur ("ce qui se passe après t "). Si X est le processus canonique, \mathcal{F}_t est la plus petite tribu qui rend mesurable toutes les "projections" $\omega \rightarrow \omega_s$ pour $s \leq t$.

Du point de vue technique, il faut supposer que la famille $(\mathcal{F}_t)_{t \in T}$ est continue à droite c'est à dire que, pour tout élément t de T ,

$$\mathcal{F}_t := \mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s$$

ceci est toujours possible en remplaçant \mathcal{F}_t par \mathcal{F}_{t+} mais peut poser un problème conceptuel.

C.7 Temps d'arrêt

Lors des **preuves heuristiques** données dans les chapitres précédents (cf. 3.E.3, le 2°) de 6.C.6, le chapitre 8) on a parfois utilisé des expressions telles que "l'instant où....". La formulation mathématique rigoureuse associée est la notion de temps d'arrêt.

Etant donné un espace fondamental (Ω, \mathcal{F}, p) muni d'une filtration $(\mathcal{F}_t)_{t \in T}$, un temps d'arrêt u est une application définie sur Ω , à valeurs

dans T , telle que, pour tout élément t de T , l'ensemble $\{\omega : u(\omega) \leq t\}$ est un élément de \mathcal{F}_t : en langage imagé, le fait de savoir si u est atteint ou non ne fait pas intervenir le futur (sauf, éventuellement, le "futur immédiat" si on a remplacé \mathcal{F}_t par \mathcal{F}_{t+}).

Cette notion de temps d'arrêt est un outil technique absolument indispensable. Elle ne peut être valablement utilisée que si les processus considérés sont continus à droite ou continus à gauche (d'où l'importance de disposer d'une modification cadlag).

Par exemple, on peut facilement (cf. le paragraphe 1.6 de [MeP]) démontrer le lemme suivant :

Lemme : Soit X un processus réel continu à droite ou continu à gauche (par trajectoires) et défini sur l'espace fondamental (Ω, \mathcal{F}, p) . Soit $(\mathcal{F}_t)_{t \in T}$ une filtration associée continue à droite. Soit u un temps d'arrêt et $a \geq 0$. On pose :

$$v(\omega) := \inf. \{t : t \in T, t > u(\omega), |X(\omega, t) - X(\omega, u(\omega))| > a\}$$

Alors v est un temps d'arrêt.

D. Constructions poissonniennes

D.1 Introduction

Le processus de Poisson possède de nombreuses propriétés tout à fait remarquables. Par ailleurs, dans certains cas particuliers, il y a une relation étroite entre la balance locale et l'existence de "flots poissonniens" (cf. 8.E). Il a donc paru intéressant de **construire les processus markoviens discrets à partir des processus de Poisson**(cf. [Bre]).

Ce paragraphe C se propose donc de donner les idées de base d'une telle construction, étant bien entendu qu'il peut y avoir d'autres variantes techniques que celles proposées ici.

D.2 Cas E fini et $T = \mathbb{R}^+$

Soit E un ensemble fini (E est l'ensemble des états possibles à chaque instant t). On suppose que l'ensemble T associé au temps est \mathbb{R}^+ . Soit a une fonction positive définie sur $(E \times E)$ telle que, pour tout élément e de E , $a(e, e) = 0$. Soit q une fonction positive définie sur E telle que

$$\sum_{e \in E} q(e) = 1.$$

On peut construire un processus $(X_t)_{t \in T}$ qui admet a comme matrice d'évolution et q comme loi initiale (loi à l'instant $t=0$) en procédant comme indiqué en B.3. On peut aussi construire le processus X à partir d'un espace fondamental "poissonnien" en procédant comme suit :

D'une part, on pose $\Omega' := E$, \mathcal{F}' est la tribu de toutes les parties de E et $p' = q$. Notons F' l'espace probabilisé $(\Omega', \mathcal{F}', p')$.

D'autre part, soit H l'ensemble des éléments (i,j) de $(E \times E)$ tels que $a(i,j) \neq 0$. A chaque élément h de H on associe un processus canonique $F_h := (\Omega_h, \mathcal{F}_h, p_h)$ associé à l'automate exponentiel (cf. B.5) de taux $u = a(i,j)$ si $h = (i,j)$ et de loi initiale stationnaire ou non.

Soit $F := (\Omega, \mathcal{F}, p)$ l'espace produit (cf., par exemple, [Met-1] de F' et des espaces $(F_h)_{h \in H}$: autrement dit, Ω est le **produit "ensembliste"** des espaces Ω' et $(\Omega_h)_{h \in H}$, \mathcal{F} est la **tribu produit** et p est la **probabilité produit**.

Cette construction mathématique peut être interprétée "physiquement" comme suit : il y a d'une part l'état initial associé à Ω' ; d'autre part, pour chaque couple $(i,j) = h$, il y a un automate exponentiel de taux $a(i,j)$ qui oscille constamment entre les états 1 et 2. Ces divers automates évoluent indépendamment les uns des autres et indépendamment de l'état initial (associé à Ω'). L'état initial de chaque automate peut être choisi quelconque.

D.3 Construction de X

On a donc, en D.2 ci-dessus, construit un espace fondamental (Ω, \mathcal{F}, p) . On va maintenant construire le processus X sur cet espace fondamental. Soit $(\mathcal{F}_t)_{t \in T}$ la filtration produit des filtrations canoniques.

Pour tout élément h de H , soit $X'(h)$ le processus associé à l'état de l'automate h ; soit $X(h)$ une modification "fixe par morceaux" et continue à droite du processus $X'(h)$ (cf. C.5).

Soit $u(n)$ la suite de variables aléatoires à valeurs dans T et définies sur Ω comme suit : $u(0) = 0$ et, pour $n \geq 0$, $u(n+1) := \inf. \{t : t > u(n), \exists h \in H, X(h)_t \neq X(h)_{u(n)}\}$.

On vérifie que la suite $(u(n))_{n \geq 0}$ est une suite croissante de **temps d'arrêt** relativement à la filtration $(\mathcal{F}_t)_{t \in T}$ et que, sauf sur un ensemble de probabilité p nulle, $\lim_{n \rightarrow \infty} u(n) = +\infty$ et les sauts des divers automates ne sont

jamais simultanés (attention : ceci est une propriété liée à l'indépendance des divers automates et non une hypothèse).

On construit alors X de la façon suivante :

En $t = 0$, $X((\omega', (\omega_h)_{h \in H}), 0) := \omega'$

Puis, pour tout couple (ω, t) tel que

$$u(n)(\omega) < t < u(n+1)(\omega), \quad X(\omega, t) = X(\omega, u(n)(\omega))$$

Enfin, pour tout couple (ω, t) tel que $t = u(n+1)(\omega)$,

$$X(\omega, u(n+1)(\omega)) = j \text{ si } X(\omega, u(n)(\omega)) = i$$

et si c'est l'automate (i, j) qui change d'état à l'instant $u(n+1)(\omega)$; sinon, on pose

$$X(\omega, u(n+1)(\omega)) = X(\omega, u(n)(\omega)) .$$

Cette construction mathématique peut être interprétée de la façon suivante : quand le processus X est dans l'état i , il effectue une transition dès que l'un des automates $(i, j)_{j \in E}$ change d'états (la transition étant la transition associée à cet automate), les autres automates (i', j) pour $i' \neq i$ n'étant pas pris en considération.

On vérifie alors, compte tenu du caractère "sans mémoire" - c'est à dire markovien - de chaque automate que le processus X admet la loi conjointe associée à la loi initiale q et à la matrice d'évolution a .

D.4 Remarques

1°) La construction qui précède admet de nombreuses variantes (cf. [Bre]). Par exemple, au lieu de considérer des automates exponentiels, on peut considérer des processus de Poisson : ceci a l'inconvénient d'introduire un ensemble d'états infini même si l'ensemble des états de X est fini mais présente certains avantages techniques.

2°) Par ailleurs, soit $(G(i))_{i \in I}$ une partition de E telle que, quel que soit i élément de I , j élément de E , k et k' éléments de $G(i)$, on a $a(k, j) = a(k', j)$. Alors on peut poser $H := (I \times E)$ au lieu de choisir $H := (E \times E)$: le même automate peut gérer toutes les transitions d'un élément quelconque de $G(i)$ à j .

3°) Certains des automates peuvent avoir une loi d'évolution non exponentielle (il n'y a rien à modifier dans la construction proposée précédemment) mais, attention, le gain de généralité ainsi obtenu est en fait très limité (cf. la fin du 2°) de D.5 ci-dessous) ; plus précisément, la base de la construction repose sur une hypothèse très restrictive d'indépendance.

D.5 Les limites de telles constructions

1°) Tout d'abord, sauf cas très particuliers, on ne peut pas prendre $T = \mathbb{R}$ (à cause de la loi initiale) : or, en théorie ergodique, il est techniquement préférable de prendre $T = \mathbb{R}$ pour que "l'opérateur de translation dans le temps" soit bijectif sur l'espace fondamental.

2°) De plus, si on veut étudier un processus en régime stationnaire, il faut choisir comme loi initiale une loi stationnaire et la construction proposée précédemment ne vaut que si les automates évoluent indépendamment de cette condition initiale : notamment, la construction proposée au paragraphe 2.5 de [Bre] ne peut être étendue au cas de la file GI/M/1 stationnaire ou, a fortiori, au cas de la file M/GI/1 stationnaire que si la loi initiale est préalablement déterminée : or cette loi initiale doit être définie sur un espace d'états très "lourd" puisqu'il doit prendre en compte conjointement l'état de la file et le délai avant la dernière arrivée ou le dernier départ.

L'auteur pense donc que, en général, il est préférable d'utiliser d'emblée la "voie royale", c'est à dire de construire un semi-groupe suffisamment général (cf. [HiP] ou [Yos] par exemple) et de construire le processus canonique associé en tant que limite projective (cf. [Met-2]).

E. Processus markovien stationnaire

E.1 Processus markovien

La construction d'une modification cadlag du processus canonique X n'est évidemment que la première étape de l'étude de X .

Quand l'ensemble E des états de X est fini, on a vu en B.3 que X est markovien si, pour toute séquence croissante $(t(k))_{1 \leq k \leq n}$ d'éléments de T et toute séquence $(e_k)_{1 \leq k \leq n}$ associée d'éléments de E , on a, pour $1 < m < n$:

$$\begin{aligned} (10E1) \quad & \text{Proba} [\forall i > m, X_{t(i)} = e_i \mid X_{t(m)} = e_m] \\ & = \text{Proba} [\forall i > m, X_{t(i)} = e_i \mid \forall j \leq m, X_{t(j)} = e_j] \end{aligned}$$

On a vu à plusieurs reprises, notamment quand on utilise la "chaîne incluse", qu'il y a lieu de considérer le processus X , non pas seulement à des instants fixés mais aussi aux "instants où..." : relativement à de "tels instants", le processus est encore "markovien". La formalisation mathématique associée consiste à dire que **la propriété 10E1 est satisfaite si, au lieu de considérer une séquence croissante d'instants "fixes", on considère une séquence croissante de "temps d'arrêt". Dans ce cas, on dit que le processus X est fortement markovien** (d'autres définitions

techniques de la propriété de Markov forte peuvent être proposées mais l'idée de base en reste la même).

Quand l'ensemble des états est fini, le processus canonique associé à un semi-groupe de transition comme défini en 1.D.5 admet une modification cadlag qui est fortement markovienne (par exemple, cf [Doo]).

La preuve de cette propriété se situe à un niveau accessible mais, **attention** : considérons par exemple, le cas où on étudie la file GI/M/1 au sens strict c'est à dire que par souci de généralisation mathématique on ne veut pas considérer une approximation par états fictifs. **Le processus X associé est markovien si l'état à un instant donné prend en compte le délai écoulé depuis la dernière arrivée : l'ensemble E des états doit donc être choisi continu.** De plus, la topologie qu'il faut mettre sur E n'est pas, en général, celle qui correspond à la topologie usuelle de \mathbb{R} .

E.2 Processus stationnaire

Pour la majorité des applications, le processus X doit être un processus markovien qui évolue **en régime stationnaire**. Il faut d'abord montrer l'existence, éventuellement l'unicité, de la "probabilité stationnaire" q - on dit aussi la "mesure invariante", etc... - par des techniques d'algèbre et d'analyse. Ensuite on construit le processus de loi q à chaque instant "fixe" (cf. B.4). Evidemment, cela ne suffit pas.

Par exemple, en 2.E.2, on a noté que le "théorème ergodique" le plus utilisé (souvent implicitement) dit que **si X est un processus stationnaire et f une fonction de l'état, on a, pour "presque toutes les trajectoires",**

$$E[f(X_t)] = \lim_{t \rightarrow \infty} \frac{1}{t} \int_s^{s+t} f(X_r) dr$$

On peut trouver une preuve de cette propriété (à temps discret et à temps continu) dans [Doo].

Sauf si on se limite à une vérification dans des cas particuliers, la preuve d'une formule aussi simple que la formule de Little (cf. 2.F.4) est "à ce prix".

E.3 Chaîne incluse

Considérons une "théorie mathématique" comprenant la construction d'un processus markovien stationnaire X et la preuve des propriétés évoquées ci-dessus : un bon critère pour tester l'efficacité de cette théorie - par rapport aux problèmes abordés dans ce cours - est de regarder si elle permet à la fois de prouver la formule de Little (cf. ci-dessus) et de justifier les calculs liés à la "chaîne incluse" (cf. 6.C).

En effet, l'étude de la chaîne incluse utilise la propriété de Markov forte - c'est à dire le caractère markovien relativement aux instants u_n des "transitions" de type H (cf. 6.C.4) - et la stationnarité. De plus, on y considère la limite du processus quand t tend vers l'infini ce qui impose, notamment de démontrer que les "temps d'arrêt" u_n sont "presque sûrement finis". Etc...

E.4 Intégrale stochastique

Considérons un processus X dont l'ensemble T des temps est compact.

Les notions qui viennent d'être évoquées ne recouvrent qu'une petite partie de la "théorie des probabilités" : notamment il est essentiel de noter que l'étude des processus markoviens quand l'ensemble des états est fini ou dénombrable - que ce soit, ou non, en régime stationnaire - **ne fait intervenir que des processus "à variation bornée par trajectoires"**. Quand l'ensemble des états est fini, sur chaque intervalle de temps borné le nombre de changements d'états est même fini "presque sûrement par trajectoires". Ceci simplifie considérablement la construction et surtout l'étude des "trajectoires".

Par contre, lorsque le problème étudié conduit à adopter une modélisation à l'aide d'un processus continu, ce processus admet, en général, "presque sûrement une variation infinie par trajectoires". Il en est ainsi pour le **mouvement brownien** et tous les processus construits à partir du mouvement brownien (**processus de diffusion**, etc...). La "formule" classique

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dX_s$$

ne vaut plus et doit être remplacée par la **formule de Ito** qui fait intervenir l'intégrale de $f'(X_s)$ par rapport à la "variation quadratique" et l'**intégrale stochastique** de $f'(X_s)$ par rapport à X .

Il s'agit alors d'une théorie complètement différente de celles évoquées précédemment (cf. [DeM] ou [MeP] par exemple).

F. Conclusion

En conclusion, on dispose de tous les outils mathématiques nécessaires pour résoudre les problèmes théoriques qui ont été évoqués dans ce cours ; toutefois, il n'y a pas d'ouvrage qui recouvre complètement le

cadre des processus markoviens discrets en régime stationnaire ; le [Doo] reste la référence de base mais des notions telles que temps d'arrêt ou filtration y sont peu formalisées.

Pour autant, est-il raisonnable de demander à tous les ingénieurs qui utilisent les techniques abordées dans ce cours d'assimiler la "grande théorie" introduite dans ce chapitre 10 ? L'auteur pense que non. L'essentiel est de savoir qu'elle existe et, lorsqu'on construit un modèle mathématique avec des hypothèses "intuitivement naturelles" de "sentir" si ces hypothèses sont, ou non, cohérentes et, en cas de doute, d'en référer à un spécialiste.

Enfin, et surtout, il ne faut pas oublier que tout le temps passé à assimiler un arsenal théorique quelque peu indigeste n'est pas utilisé pour améliorer les techniques proposées dans ce cours : or, dans ce domaine, l'essentiel reste à faire.

REFERENCES

- [Abd] H. ABDALLAH, *Construction d'un logiciel de calcul des éléments transitoires de chaînes de Markov à temps continu*, Thèse, Rennes, 1989 : 8.A
- [Alg] ALGERENNES, *Probabilités stationnaires pour des systèmes markoviens discrets*, rapport INRIA n° 633, 1987 : 5.D.4, 5.E.3
- [Ast] M.A. ASTOUATI, *Méthode des convexes pour le calcul de la probabilité stationnaire d'un réseau à deux stations à lois de services exponentielles*, Thèse de magister, Alger, 1985 : 5.D.4
- [BCMP] F. BASKETT, K.M. CHANDY, R.R. MUNTZ and F.G. PALACIOS, *Open, closed and mixed networks of queues with different classes of customers*, J.A.C.M. vol. 22, n° 2, 248-260, 1975 : 4.C
- [BDK] P. BOYER, A. DUPUY et A. KHELLADI, *Méthodes de résolution d'un système linéaire markovien*, Preprint, 1987 : 5.D.4
- [Bec] M. et P. BECKER, *Importance des erreurs dues à la précision limitée de l'ordinateur pour la validation des simulations*, RAIRO Informatique, n° 4, 1977 : 3.F
- [Bel] R. BELLAMINE, *M-matrices et probabilités stationnaires*, Thèse, Rennes, 1989 : 5.E
- [BeP] R. BELLAMINE et J. PELLAUMAIL, *Opérateurs positifs et probabilités stationnaires*, RAIRO, Recherche opérationnelle, n° 3, 1989 : 5.E
- [Ber] C. BERGE, *Théorie des graphes et ses applications*, Dunod, 1967 : 7
- [Bha] A.T. BHARUCHA-REID, *Elements of the theory of Markov processes and their applications*, McGraw Hill, 1960 : 9.C.4
- [BhW] BHANDIWAD et A.C. WILLIAMS, *A generating function approach to queueing network analysis of multiprogrammed computers*, Networks 6,1,1-22, 1976 : 4.B.4
- [Bir] G. BIRKHOFF, *Extension of Jentzsch's theorem*, Trans. Amer. Math. Soc. 85, 219-227, 1957.
- [BKP] R. BELLAMINE, A. KHELLADI et J. PELLAUMAIL, *Opérateurs positifs, fonctions génératrices et probabilités stationnaires*, Rapport L.A.N.S. n° 10, Rennes, 1988 : 7.C.3
- [BoM] R. BOTT et J.P. MAYBERRY, *Matrices and trees*, Economic activity analysis, Wiley, 1954 : 7.C.3

- [BoP] P. BOYER et J. PELLAUMAIL, *Deux files d'attente à capacité limitée en tandem*, Rapport IRISA n° 147, 1981 : 5.D.4
- [Boy] P. BOYER, *Evaluation des performances des systèmes réparés*, Thèse, Rennes, 1982 : 5.D.4
- [BPe] A. BERMAN and R.J. PLEMMONS, *Nonnegative matrices in the mathematical sciences*, Academic Press, 1979 : 5,6,7.
- [BrB] S.C. BRUELL et G. BALBO, *Computational algorithms for closed queueing networks*, North-Holland, 1980 : 4.B.4
- [Bre] P. BREMAUD, *Point processes and queues*, Springer-Verlag, 1981 : 8.C.4, 10.D
- [Buz] J.P. BUZEN, *Fundamental operational laws of computer performance*, Acta informatica 7, 167-182, 1976 : 4.B.4
- [Car] H. CARTAN, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann, 1961 : 9
- [Ch1] F. CHARLOT, *Systèmes de files d'attente*, Thèse, Rouen, 1988 : 10.B.6
- [Ch2] F. CHATELIN, 5.4.2
1 – *Valeurs propres de matrices*, Masson, 1988.
2 – *Analyse statistique de la qualité numérique et arithmétique de la résolution approchée d'équations par le calcul sur ordinateur*, Rapport IBM, FI33, 1988.
- [Che] W.K. CHEN, *Applied graph theory*, North-Holland, 1971 : 7.C.3
- [ChF] P. CHRETIENNE et R. FAURE, *Processus stochastiques, leurs graphes, leurs usages*, Dunod, 1974.
- [ChT] M.L. CHAULDRY et J.G.C. TEMPLETON, *A first course in bulk queues*, Wiley, 1983.
- [Chu] K.L. CHUNG, *Markov chains with stationary transition probabilities*, Springer-Verlag, 1967 : 5.C.4
- [Cia] P.G. CIARLET, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, 1982 : 5.A.2
- [Cin] E. CINLAR, *Introduction to stochastic processes*, Prentice-Hall, 1975 : 8.E., 10.B.4
- [Coh] COHEN, *The single server queue*, North-Holland, 1969.
- [Cou] P.J. COURTOIS, *Decomposability*, Academic Press, 1977.

- [Cox] D.R. COX, *A use of Complex probabilities in the theory of stochastic processes*, Proc. Cambridge Philosophical Society, vol. 51, p. 313-319, 1955 : 1.F.5
- [DeM] C. DELLACHERIE et P.A. MEYER, *Probabilités et potentiel*, Hermann, 1980 : 10.E.4
- [DER] I.S. DUFF, A.M. ERISMAN et J.K. REID, *Direct methods for sparse matrices*, Clarendon Press, 1986 : 5.A.3
- [Des] G. DESBAZELLE, *Exercices et problèmes de recherche opérationnelle*, Dunod, 1972.
- [Doo] J.L. DOOB, *Stochastic processes*, Wiley, 1953 : 8.E, 10.E
- [Dud] L. DUDKIN, *Iterative aggregation techniques in soviet national economic planning*, Monograph series on soviet union, Delphic-associates, 1987 : 6.F
- [Fau] F. FAURE, *Précis de recherche opérationnelle*, Dunod, 1971 : 1.C.3
- [Fel] W. FELLER, *An introduction to probability theory and its applications*, Wiley, 1965.
- [Gan] F.R. GANTMACHER, *Théorie des matrices*, Dunod, 1966 : 2.B.3
- [GeP] E. GELENBE et G. PUJOLLE, *Introduction aux réseaux de files d'attente*, Eyrolles, 1982.
- [GLM] E. GELENBE, J. LABETOULLE, R. MARIE, etc..., *Réseaux de files d'attente*, AFCET Hommes et techniques, 1980.
- [GoM] G.H. GOLUB et G.A. MEURANT, *Résolution numérique des grands systèmes linéaires*, Eyrolles, 1983 : 5.A.2, 6.A.3
- [GoN] W.J. GORDON et G.F. NEWELL, *Closed queueing systems with exponential servers*, Operations Research, 15, pp. 254-265, 1967 : 4.B.2
- [GoV] G.H. GOLUB et C.F. Van LOAN, *Matrix computations*, The Johns Hopkins University press, 1983 : 5.A.2
- [GrH] D. GROSS et C.M. HARRIS, *Fundamentals of queueing theory*, Wiley, 1974.
- [GuP] R. GUIHUR et J. PELLAUMAIL, *An analytical study of packet satellite multiple access*, Computer Networks 7, pp. 389-393, 1983.
- [Heb] G. HEBUTERNE, *Ecoulement du trafic dans les autocommutateurs*, Masson, 1985 : 3.A.3

- [HeP] J.M. HELARY et R. PEDRONO, *Recherche opérationnelle*, Travaux dirigés, Hermann, 1983 : 1.C.3
- [HiP] E. HILLE and R.S. PHILIPPS, *Functional analysis and semi-groups*, American Mathematical society, New York, 1957 : 10.D.5
- [InR] P. INGELS et M. RAYNAL, *Simulation répartie de systèmes à évènements discrets, partie 1*, Rapport INRIA n° 1057, 1989.
- [Jac] J.R. JACKSON, *Jobshop-like queue system*, Management Sci., vol. 10, pp. 131-142, 1963 : 4.B.2
- [KeI] F.P. KELLY, *Resersibility and stochastic networks*, Wiley, 1979 : 5.E.3 ; 8.C.4, 8.D
- [Khl] A. KHELLADI, *Changement de variable et probabilités stationnaires*, Preprint. : 5.D.4
- [Khm] Z. KHEMSI, *File d'attente à plusieurs classes de clients et arrivées non exponentielles*, Thèse de magister, Alger, 1989 : 5.D.4
- [Kle] L. KLEINROCK, *Queueing systems*, Wiley, 1976.
- [Lag] O. LAGHA, *Approximation dans les réseaux de files d'attente*, Thèse de magister, Alger, 1989 : 7
- [Lar] T. LARDJANE, *Conditions suffisantes de réversibilité et de composition de réseaux de files d'attente*, Thèse de magister, Alger, 1985 : 5.E.3
- [L1b] M. LEBAH, *Probabilités stationnaires pour des réseaux à deux stations et à plusieurs classes de clients*, Thèse de magister, Alger, 1985 : 4.A.7
- [L1P] M. LEBAH et J. PELLAUMAIL, *Balance locale dans les réseaux à trois stations*, Ann. Scient. Univ. Clermont-Ferrand, 1985 : 4.A.7
- [L2b] J.Y. LE BOUDEC, *Analyse quantitative de réseaux de files d'attente markoviens*, Thèse, Rennes, 1984 : 4.A.7, 8.C.4
- [L2P] J.Y. LE BOUDEC et J. PELLAUMAIL, *Equations de Chapman-Kolmogorov et flots stationnaires pour des processus markoviens*, Rapport INRIA n° 260, 1983 : 8.C.4
- [Lem] B. LEMAIRE, 5.E.3, 8.F.6
 - 1 – *Méthode de conservation et blocage dans les files d'attente*, RAIRO, Recherche Opérationnelle, vol. 11, n°4, pp. 363-377, 1977.
 - 2 – *Une démonstration directe de la formule de Pollaczek-Khintchine*, RAIRO, Recherche Opérationnelle, vol. 12, n° 2, pp. 229-231, 1978.
 - 3 – *Théorème de conservation des clients dans les files d'attente*, RAIRO, Recherche Opérationnelle, vol. 12, n° 4, pp. 395-399, 1978.
 - 4 – *Régularité des matrices à diagonale dominantes. Applications à l'absorption dans les chaines et processus de Markov*, RAIRO,

Recherche Opérationnelle, vol. 19, n° 3, 1985.

- [Len] L.M. LE NY,
1 – *Etude analytique de réseaux de files d'attente multi-classes à routages variables*, RAIRO, vol. 14, n° 4, pp. 331-347, 1980 : 4
2 – *Probabilités stationnaires d'une file d'attente à deux seuils*, Preprint.
- [Leo] W.W. LEONTIEF, *The structure of american economy*, IASP, 1951 : 6.F
- [Ler] J. LEROUDIER, *La simulation à événements discrets*, AFCET, Hommes et Techniques, 1980 : 3.F.1
- [LeW] L.M. LE NY et C.M. WOODSIDE, *Performance modelling with rendez-vous service*, Rapport INRIA n° 941, 1988.
- [LiN] Z. LIU et P. NAIN, *Sensitivity results in open, closed and mixed product-form queueing networks*, Rapport IRISA n° 1144, décembre 1989 : 4.B.5
- [LPS] J. LABETOULLE, G. PUJOLLE et C. SOULA, *Stationnary distributions of flows in Jackson networks*, Mathematics of operations research, vol. 6, n° 2, 1981 : 8.C.4
- [MaP] R. MARIE et J. PELLAUMAIL, *Steady-state probabilities for a queue with a general service distribution and state-dependent arrivals*, IEEE, vol. SE-9, n° 1, 1983 : 5.E.5
- [Mar] R. MARIE, 1 – *Modélisation par réseaux de files d'attente*, Thèse, Rennes, 1978 : 5.E.3
- [Mat] N.S. MATLOFF, *Probability modelling and computer simulation*, PWS-KENT, 1988.
- [MeP] M. METIVIER et J. PELLAUMAIL, *Stochastic integration*, Academic Press, 1980 : 10.C, 10.E.4
- [Met] M. METIVIER,
1 – *Notions fondamentales de la théorie des probabilités*, Dunod, 1972.
2 – *Limites projectives de mesures. Martingales. Applications*. Annali di Matematica pura ed applicata (IV), vol. LXIII, pp. 225-352, 1963 : 10.D.5
- [MRT] R. MARIE, A.L. REIBMAN et K.S. TRIVEDI, *Transient analysis of acyclic Markov chains*, Performance evaluation 7, pp. 175-194, 1987 : 8.A.7
- [Neu] M.F. NEUTS, *Matrix-geometric solutions in stochastic models*, The Johns Hopkins University Press, London, 1981 : 5.C.4, 7.D.5
- [Nev] J. NEVEU, *Processus ponctuels*, Ecole d'été de probabilités de Saint-Flour, Lecture Notes n° 598, Springer-Verlag, 1976 : 10.B.6

- [Oum] B. OUMEHDI, *Modélisation avec pivot d'un réseau de files d'attente à lois d'arrivée et de départ non exponentielles*, Thèse de magister, 1985 : 5.D.4
- [OPS] B. OUMEHDI, J. PELLAUMAIL et H.SAGGOU, *Matrices creuses et probabilités stationnaires*, Rapport LANS, juin 1990 : 5.D.5
- [PeA] H.G. PERROS et T. ALTIOK, *Queueing networks with blocking*, North-Holland, 1989.
- [Pel] J. PELLAUMAIL,
 - 1 – *Sur un problème de fiabilité*, Rapport, Rennes, 1967 : 8.B.6
 - 2 – *Formule du produit et décomposition de réseaux de files d'attente*, Ann. Inst. Henri Poincaré, vol. XV, n° 3, pp. 261-2864, 1979 : 4
 - 3 – *Disponibilité d'un ensemble complexe*, RAIRO, Recherche Opérationnelle, vol. 14, n° 2, 1980 : 2.C.2
 - 4 – *Solution à forme produit d'un système linéaire*, Rapport INRIA, n° 259, 1983 : 4
 - 5 – *Modélisation avec pivot pour une loi générale*, RAIRO, Recherche Opérationnelle, vol. 19, n° 3, 1985 : 1.F.4, 6.B
 - 6 – *Probabilités, statistiques, files d'attente*, Dunod, 1986.
 - 7 – *Graphes et probabilités stationnaires*, Ann. Inst. Henri Poincaré, vol. 26, n° 1, 1990 : 2.A.7
- [Pis] S. PISSANETZKY, *Sparse matrix technology*, Academic Press, 1984 : 5.A.3
- [PSS] B. PHILIPPE, Y. SAAD et W.J. STEWART, *Numerical methods in Markov chain modeling*, Rapport IRISA n° 495, 1989 : 5.A.2
- [Rev] D. REVUZ, *Markov chains*, North-Holland, 1975 : 2.B.3, 6.A.3, 7.D.6
- [Ros] ROSEAUX, *Exercices et problèmes résolus de recherche opérationnelle*, Masson, 1987 : 1.C.3
- [Saa] T.L. SAATY, *Elements of queueing theory*, Mac Graw-Hill, 1961.
- [Sag] H. SAGGOU, *Etude d'une file avec priorité et lois non exponentielles*, Thèse de magister, Alger, 1989 : 5.C.3, 5.D.4, 7.A.1
- [Sen] E. SENETA, *Non-negative matrices and Markov chains*, Springer-Verlag, 1981 : 2.A.3, 2.B.3, 6.F
- [SiM] M. SIBONY et J.C. MARDON, *Approximations et équations différentielles*, Hermann, 1982.
- [Ste] G.W. STEWART, *Introduction to matrix computation*, Academic Press, 1973.
- [Sys] R. SYSKI, *Introduction to congestion theory in telephone systems*, North-Holland, 1986.

- [Tho] W.A. THOMPSON, *Point process models with applications to safety and reliability*, Chapman and Hall, 1988.
- [Val] G. VALIRON, *Equations fonctionnelles, Applications*, Masson, 1950 : 9
- [Var] R.S. VARGA, *Matrix iterative analysis*, Prentice Hall, 1962 : 10.A.2
- [Wal] V. WALLACE, *The solution of quasi birth and death processes arising from multiple access computer systems*, PH. D. diss., Systems Engineering Laboratory, University of Michigan, Techn. report n° 07742-6-T, 1969 : 7.D.5
- [Yos] YOSIDA, *Functional analysis*, Springer Verlag, 1971 : 10.D.5

INDEX

- A** absorbant (point) 6.D – acyclique 5.A.1 – angulaire (écart) 2.A.2 – anneau 7.A.3 – arrêt (temps d') 3.E.4, 10.C.7 – attente (temps d') 8.F.3 – automate (exponentiel) 1.E.6, 10.B.5 – axiome du choix 10.A.3.
- B** balance locale 4.A.5, 4.B.2, 4.C, 4.D.4, 4.E.4, 4.F, 8.C.3 – BCMP (théorème) 4.C.2, 4.F – binomiale (loi) 1.A.2, 9.A.2 – blocs (notation par) 5.A.1 – Bochner (théorème de Kolmogorov...) 10.B.2 – buffer 5.E.5 – bus 3.D.3.
- C** cadlag 10.C – canonique (processus) 10.B – caractéristiques (méthode des) 9.C.5, 9.E.2 – carrée (matrice) 5.A.1 – central (serveur) 4.A. – chaîne (de Markov) 1.C, 2.C.3, 3.D – chaîne incluse 6.C, 8.B, 8.F.4, 8.F.6, 9.B.3, 10.E.3 – Chapman(Kolmogorov) 1.D.2 – codage 3.B – contraction (coefficient de) 2.A – convexes (méthodes des) 5.D – Cox 1.F.5.
- D** définition 7.A.2 – déterminant 7.C.3 – diagonale (matrice) 5.A.1 – différentiel (système...linéaire) 1.D.4,8 – discrétisation (du temps) 3.E.2 – durée de vie 8.F.3.
- E** entière (variable aléatoire) 9.A.2 – équilibre (équations d') 2.D – ergodique 2.D.5, 2.E – Erlang (loi d') 1.E.2, 1.E.5, espace fondamental 10.A.5 – évolution (matrice d') 1.D.1 – explosion 2.D.5 – exponentielle (loi) 1.A.4, 1.E.1 – exponentielle (de matrice) 8.A.
- F** factorisation (d'une matrice) 5.B.2 – fiabilité 6.D.3, 8.F – fictif (état) 1.F – fifo 4.F.2 – files (d'attente) 4 – file (unique) 1.E.4, 1.E.5, 3.E.3, 5.E.5, 8.F.6, 9.B, 9.E – filtration 10.C.6 – fluvial (réseau) 7.A – forme (produit) 4 – Frobenius (théorie de Perron...) 2.B.3.
- G** générateur (de nombres au hasard) 3.B – génératrice (fonction) 9 – graphe 7.
- H** hasard (nombres au) 3.B – heuristique (preuve) 10.A.2 – hiérarchique (dispatching) 4.A.7 – homogène 1.A.4, 1.B.4 – hyperexponentielle 1.F.3.
- I** indistingables (processus) 10.C.3 – infini (cas E...) 1.D.4, 3.A.6, 5.E.6, 7.D.6 – inversion (d'une matrice) 6.A.4, 6.C.2 – irrigabilité 7.A.5 – irréductible (matrice) 1.B.3, 2.D.3, 5.A.2, 7.A.1.
- J** Jackson (théorème de) 4.B.2 – jeu 1.C, 3.D.2.
- K** Kolmogorov (équations de) 1.D.2, 1.D.3 – Kolmogorov (théorème de... Bochner) 10.B.2.

- L** Langage (de simulation) 3.F.1 – Léontief (matrices de) 6.F – Little (formule de) 2.F.4 – L-matrice 5.A.1 – loi (d'un processus) 1.B.
- M** macro-état 4.C.1 – marche aléatoire 1.A.2 – Markov, markovien 1, 2.F.3, 3.A.2, 3.A.4, 3.E, 10.E.1 – matrice 5 – mélange 2.D.5 – mer 7.A.3 – micro-état 4.C.1 – M-matrice 5.A.1, 7.A.6 – modélisation 3.C.3 – modification (d'un processus) 10.C.2 – monotone (matrice) 5.A.1 – Monte-Carlo (méthode de) 3.D.6 – morts (processus de) 9.C.
- N** naissances (processus de) 9.C – nombres (au hasard) 3.B – non-explosion 2.D.5 – normalisation (constante de) 4.B.4.
- P** partielles (équations aux dérivées) 6.E, 9 – pénalisation 6.D.4 – périodique (matrice) 2.C.3 – Perron (théorie de ...Frobenius) 2.B.3 – PH (loi) 1.F.2 – pivot 1.F.4, 6.B – poissonnienne (construction) 10.D – poissonnien (flot) 8.C – Poisson (loi de) 1.A.3, 9.A.2 – Poisson, poissonnien (processus) 1.A.3, 1.E.3, 3.A.3, 8.C, 10.B.6 – Pollaczek (formule de ...Khinchine) 9.B.2 – positives (produit de matrices) 2.B – précision 7.C.2 – premier arrivé, premier servi 4.F.2 – primitive (matrice) 2.C.3 – processus 1.B – produit (de réseaux) 4.E – produit (forme) 4 – projectif (système) 1.B.2, 10.B.2 – propre (réseau fluvial) 7.A.3 – propre (vecteur) 2.B.1.
- Q** QNAP2 3.F.
- R** renouvellement (processus de) 8.E – renversement (du temps) 8.D – répétition (d'appels) 3.A.4 – réversible (processus) 8.D – rigueur mathématique 10.A.2 – routages (fixes) 4.A.6, 4.B, 4.C, 9.D, 9.F – routages (variables) 4.A.7.
- S** séjour (temps de) 8.F.3 – semi-groupe (de transition) 1.B.4, 1.C.2, 1.D.5 – serveur (central) 4.A – simulation 3 – sondage 3.B – splitting 6.A, 8.B – stabilité (théorème de) 4.D.3 – stationnaire 2, 3.A.5, 3.D, 3.E, 4.A.4, 9.B, 10.B.4, 10.E.2 – stochastique (matrice) 1.C.3, 2.A.5, stochastique (processus) 1.B.1 – stock 1.C.6 – sureté (de fonctionnement) 6.D.3, 8.F.
- T** temps (d'arrêt) 3.E.4, 10.C.7 – trajectoire 2.E, 2.F – transition (semi-groupe de) 1.B.4, 1.C.2, 1.D.5 – transitoire 8, 9 – translation (opérateur de) 1.A.4 – triangulaire 8.B.6 – tridiagonale (matrice) 5.B, 5.C, 6.A.5, 6.B.4, 7.D.

LISTE DES DERNIERES PUBLICATIONS INTERNES IRISA

1991

- PI 570: DESIGN DECISION FOR THE FTM : A GENERAL PURPOSE FAULT TOLERANT MACHINE**
Michel BANATRE, Gilles MULLER, Bruno ROCHAT, Patrick SANCHEZ
Janvier 1991, 30 pages
- PI 571 ANIMATION CONTROLEE PAR LA DYNAMIQUE**
Georges DUMONT, Marie-Paule GASCUEL, Anne VERROUST
Février 1991, 84 pages
- PI 572 MULTIGRID MOTION ESTIMATION ON PYRAMIDAL REPRESENTATIONS FOR IMAGE SEQUENCE CODING**
Nadia BAAZIZ, Claude LABIT
Février 1991, 48 pages
- PI 573 A SURVEY OF TREE-TRANDUCTIONS**
Jean-Claude RAOULT
Février 1991, 18 pages
- PI 574 THE OPTIMAL ADAPTIVE CONTROL USING RECURSIVE IDENTIFICATION**
Anatolij B. JUDITSKY
Février 1991 - 26 pages
- PI 575 MANUEL SIGNAL**
Patricia BOURNAI, Bruno CHERON, Bernard HOUSSAIS, Paul LE
Paul LE GUERNIC
Février 1991, 84 pages
- PI 576 AN INFORMATION BASED RELIABILITY PREDICTOR FOR SYSTEMS IN OPERATIONAL PHASE**
Kamel SISMAIL
Février 1991 - 22 pages
- PI 577 MULTISCALE STATISTICAL SIGNAL PROCESSING AND RANDOM FIELDS ON HOMOGENEOUS TREES**
Albert BENVENISTE, Michèle BASSEVILLE, Ramine NIKOUKHAH,
Alan S. WILLSKY, Ken C. CHOU
Mars 1991 - 18 pages
- PI 578 TOWARDS A DECLARATIVE METHOD FOR 3D SCENE SKETCH MODELING**
Stéphane DONIKIAN, Gérard HEGRON
Mars 1991 - 22 pages
- PI 579 SYSTEMES MARKOVIENS DISCRETS STATIONNAIRES ET APPLICATIONS**
Jean PELLAUMAIL
Mars 1991 - 284 pages

Imprimé en France

par

. l'Institut National de Recherche en Informatique et en Automatique

ISSN 0249 - 6399